

A/\$

# PATENT APPLICATION TRANSMITTAL LETTER

Attorney Docket No.: A-433B

To the Assistant Commissioner for Patents:

In re the patent application of: **Lea Anne Harrington and Murray O. Robinson**  
For: **NOVEL GENES ENCODING TELOMERASE PROTEINS**

Transmitted herewith are:

- ☒ 97 pages of specification, 0 pages of sequence listing, 4 pages of claim(s) and 1 pages of abstract, totaling 102 pages.
- ☒ an **unsigned** declaration by the applicant(s).
- ☒ 46 sheets of drawing.
- ☐ a certified copy of \_\_\_\_\_
- ☐ Other: \_\_\_\_\_
- ☒ Please amend the specification by inserting before the first line the following:

- **Cross-References**

This application is a ☒ continuation-in-part ☐ continuation ☐ division of  
U.S. Serial No. 08/873,039 filed June 11, 1997, which is a continuation-in-part of U.S. Serial No. 08/751,189, filed November 15, 1996.--

- ☐ Please amend the specification by inserting before the first line the following:  
--This is a nonprovisional Application derived from U.S. Provisional Application Serial No. \_\_\_\_\_, filed \_\_\_\_\_.--

## CLAIMS AS FILED

For	Number Filed		Number Extra		Rate		Fee
Total Claims	32	- 20 =	12	x	\$22.00	=	\$ 264.00
Independent Claims	14	- 3 =	11	x	\$82.00	=	902.00
Multiple Dependent Claims	0			+	\$270.00	=	0.00
Basic Fee					\$790.00	=	790.00
						Total Filing Fee	\$1,956.00

- ☒ Please charge Deposit Account No. 01-0519 in the name of Amgen Inc. in the amount of **\$1,956.00**. **An original and one copy are enclosed.**
- ☒ The Commissioner is hereby authorized to charge any additional filing fees which may be required by the accompanying application, any additional fees which may be required during pendency of this application or credit any overpayment to Deposit Account No. 01-0519.

Respectfully submitted,

*Nancy A. Oleski*

Nancy A. Oleski  
Attorney/Agent for Applicant(s)  
Registration No.: 34,688  
Phone: (805) 447-6504  
Date: October 15, 1997

Please send all future correspondence to:  
U.S. Patent Operations/NAO  
M/S 10-1-B  
AMGEN INC.  
Amgen Center  
1840 De Havilland Drive  
Thousand Oaks, California 91320-1789, U.S.A.

**DECLARATION AND POWER OF ATTORNEY**

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first, and sole inventor (if only one name is listed below) or a joint inventor (if plural names are listed below) of the invention entitled

**NOVEL GENES ENCODING TELOMERASE PROTEINS**

which is described and claimed in the specification which:

- ☒ is attached hereto.  
☐ was filed on \_\_\_\_\_  
 as Application Serial No. \_\_\_\_\_  
 and was amended \_\_\_\_\_ (if applicable)  
 on \_\_\_\_\_

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, § 1.56(a).

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, § 112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, § 1.56(a) which occurred between the filing date of the prior application and the filing date of this application:

APPLICATION SERIAL NO.	FILING DATE	STATUS
08/873,039	June 11, 1997	Pending
08/751,189	November 15, 1996	Pending

Power of Attorney: As a named inventor, I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application and transact all business in the Patent and Trademark Office connected therewith:

Nancy A. Oleski, Registration No.: 34,688; Ron K. Levy, Registration No.: 31,539; and Steven M. Odre, Registration No.: 29,094 said attorney(s)/agent(s) to have in addition full power of revocation, including the power to revoke any power herein granted.

**DECLARATION AND POWER OF ATTORNEY (cont'd)**

Please send all future correspondence to:

U.S. Patent Operations/NAO  
 M/S 10-1-B  
 AMGEN INC.  
 Amgen Center  
 1840 De Havilland Drive  
 Thousand Oaks, California 91320-1789

Direct Telephone Calls To:

Nancy A. Oleski  
 Attorney/Agent for Applicant(s)  
 Registration No.: 34,688  
 Phone: (805) 447-6504  
 Date: October 15, 1997

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

Full Name of Sole  
 or First Inventor:

Lea Anne Harrington

Inventor's Signature:

Date:

Residence and  
 Post Office Address:

55 Pears Avenue, Toronto, Ontario, M5R 1S9, Canada

(Address, City, State, Zip Code, Country)

Citizenship:

Canadian

Full Name of Second  
 Joint Inventor, if Any:

Murray O. Robinson

Inventor's Signature:

Date:

Residence and  
 Post Office Address:

22623 Pacific Coast Highway, Malibu, California 90265, U.S.A.

(Address, City, State, Zip Code, Country)

Citizenship:

U. S. A.

## NOVEL GENES ENCODING TELOMERASE PROTEINS

This application is a continuation-in-part of  
5 United States Serial Number 08/873,039 filed 11 June  
1997, which is a continuation-in-part of United States  
Serial Number 08/751,189 filed 15 November 1996.

10 Field of the Invention

This invention relates to novel genes encoding  
polypeptides that comprise components of the telomerase  
enzyme complex, as well as to methods of making and  
using the genes and polypeptides, and assays for  
15 detecting telomerase activity.

## BACKGROUND

Related Art

20 Many physiological changes occur as humans  
age. In addition to those observed at the phenotypic  
level such as change in hair color, appearance of skin,  
decreased lean body mass, etc., there are many changes  
at the cellular and biochemical levels. One such change  
25 that has been observed is a marked decrease in the  
length of telomeres in somatic cells as they age (Harley  
*et al.*, *Nature*, 345:458-460 [1990]). Telomeres are  
repetitive DNA sequences that are localized to the ends  
of every chromosome, and are necessary for proper  
30 chromosome maintenance, replication, and localization of  
the chromosomes within the cell nucleus. Certain  
proteins such as "telomeric repeat binding factor", or  
"TRF" have been shown to interact with the telomeric DNA  
sequences (Chong *et al.*, *Science*, 270:1663-1667 [1995]).



In most organisms, telomeres are synthesized and maintained by an enzyme known as telomerase. Telomerase is a ribonucleoprotein composed of RNA and protein components, and both types of components are  
5 necessary for activity (see for example, Greider, *Annu. Rev. Biochem.*, 65:337-365 [1996]; Greider et al., in *Cellular Aging and Cell Death*, Wiley-Liss Inc., New York, NY, pp. 123-138 [1996]).

Most cells of adult humans do not have  
10 telomerase activity; exceptions include, for example, germline tissues (sperm cells and oocytes) and certain blood cells (Greider et al., *Cellular Aging and Cell Death*, supra). Decreased telomere length correlates well with decreased replicative capacity of cells in  
15 culture (referred to as cellular senescence or cell age). It has been postulated that shortened telomeres may be involved in the inability of cells to continue dividing (Harley, supra; Levy et al., *J. Mol. Biol.*, 225:951-960 [1992]; and Harley et al., *Cold Spring*  
20 *Harbor Symposium on Quantitative Biology*, 59:307-315 [1994]), thereby contributing to senescence of the cells.

Although the molecular details of the mechanism by which telomere length decreases with each  
25 successive cell division are not clear, recent reports propose various models. For example, Marcand et al. (*Science*, 275:986-990 [1997]; see also Barinaga, *Science*, 275:928 [1997]) describe a "protein counting mechanism" in yeast in which the amount of the protein  
30 RapI bound to the telomeres purportedly affects telomere length. In separate studies, van Steensel et al and Cooper et al. (*Nature*, 385:740-743 [1997]; *Nature*, 385:744-747 [1997], respectively; see also Shore, *Nature*, 385:676-677 [1997]) demonstrate that the  
35 telomere repeat binding factor, TRF, purportedly affects telomere elongation in yeast and in humans.

Recently, it has been shown that the telomeres of one class of white blood cells, called CD28-/CD8+ T-cells, are significantly shorter in AIDS patients as compared with the same cells obtained from healthy persons of the same or similar age (Effros *et al.*, *AIDS*, 10:17-22 [1996]).

In many human cancerous cells, it has been shown that telomere length does not decrease, and telomerase activity is present, regardless of the age of these cells (Kim *et al.*, *Science*, 266:2011-2015 [1994]; and Counter *et al.*, *EMBO J.*, 11:1921-1929 [1992]). It has been suggested that inhibition of telomerase in cancer cells might serve to decrease the proliferation of these cells (Harley *et al.*, *Cold Spring Harbor Symposium on Quantitative Biology*, *supra*; and Greider *et al.*, *Cellular Aging and Cell Death*, *supra*).

The RNA component of telomerase in several mammals has been cloned and sequenced (see PCT patent application WO 96/01835, published 25 January 1995; Blasco *et al.*, *Science*, 269:1267-1270 [1995]; Feng *et al.*, *Science*, 269:1236-1241 [1995]), and it has been demonstrated that this RNA component is necessary for telomerase activity (Blasco *et al.*, *supra*; Feng *et al.*, *supra*; oral presentations at Cold Spring Harbor Laboratory Conference on Telomeres and Telomerase, 3-6 November 1996). In mouse tumor models, an increase in telomerase RNA correlates with increased tumor progression (Blasco *et al.*, *Nature Genetics*, 12:200-204 [1996]). However, Avilion *et al.* (*Cancer Res.*, 56:645-650 [1996]) showed that the presence of telomerase RNA in various human tumor tissues and cell lines was not a good predictor of the presence or amount of telomerase activity in these tissues and cell lines.

Recently, Blasco *et al.* (*Cell*, 91:25-34 [1997]; see also Zakian, *Cell*, 91:1-3 [1997]) generated mice deficient for the telomerase RNA gene. Cells of these

mice apparently lack telomerase activity, but purportedly can be immortalized in culture and are able to generate tumors.

A recent report by Kirk *et al* (*Science*, 275:1478-1481 [1997]; see also Hawley, *Science*, 275:1441-1442 [1997]) describes preparation of a telomerase RNA molecule mutant that, in *Tetrahymena*, purportedly alters the ability of germline nuclei to separate during cell division.

In ciliates (single celled eukaryotic organisms), it has been found that the protein portion of telomerase is comprised of two distinct polypeptides, termed p80 and p95 (see PCT patent application WO 96/19580, published 27 June 1995; Harrington *et al.*, *J. Biol. Chem*, 270:8893-8901 [1995]; and Collins *et al.*, *Cell*, 81:677-686 [1995]). Recently, two telomerase polypeptides of molecular weight 123 kDa and 43 kDa have reportedly been purified in *Euplotes*, a single-celled eukaryotic organism (Lingner *et al.*, *Proc. Natl. Acad. Sci. USA*, 93:10712-10717 [1996]). The 123 kDa protein, for which a yeast homolog (termed "EST2") has now been identified, purportedly has reverse transcriptase motifs (Lendvey *et al.*, *Genetics*, 144:1399-1412 [1996]; Lingner *et al.*, *Science*, 276:561-567 [1997]; see also Barinaga, *Science*, 276:528-529 [1997]). Reverse transcriptase motifs such as those described by Xiong *et al* (*EMBO J.*, 9:3353-3362 [1990]) are known to be important for functional reverse transcriptase enzymatic activity. Certain mutants of this yeast homolog protein purportedly have decreased telomerase activity (Counter *et al.*, *Proc. Natl. Acad. Sci USA*, 94:9202-9207 [1997]).

A recent nucleic acid sequence entry in the Washington University/NCI Human EST Project Database, accession number AA281296, has some sequence similarity to both the *Euplotes* 123 kDa protein and the yeast homolog thereof. Two recent publications describe the

cloning of a human gene that purportedly encodes the catalytic subunit of telomerase (Nakamura et al., *Science*, 277:955-959 [1997]; Meyerson et al., *Cell* 90:785-795 [1997]).

5 Prior to the present invention, the protein component or components of mammalian telomerase had not been identified.

10 Recently, a 347 base pair nucleic acid molecule was deposited in the public database Genbank as accession number H33937. This nucleic acid molecule was apparently identified from rat PC-12 cells that had been treated with NGF (neurotrophic growth factor). No function for this nucleic acid molecule or the protein encoded by it is set forth in the Genbank database  
15 information, however, a portion of this molecule has been found to be highly homologous to a region of the mouse telomerase RNA interacting protein 1 (TRIP1) of the present invention. The polypeptide sequence of human TRIP1 has recently been identified (Harrington et al., *Science*, 275:973-977 [1997]; Nakayama et al., *Cell*,  
20 88:875-884 [1997]).

In view of the devastating effects of cancer and AIDS, there is a need in the art to identify  
25 molecules in the human body which may have an important role in the etiology of these diseases, and to manipulate the expression of such molecules in patients suffering from these and related diseases.

Accordingly, it is an object of this invention  
30 to provide nucleic acid molecules and polypeptides that are components of the telomerase enzyme complex and can affect aging and/or proliferation of cells in the human body.

It is a further object to provide methods of  
35 altering the level of expression of such polypeptides in the human body.

Other related objects will readily be apparent from a reading of this disclosure.

#### SUMMARY OF THE INVENTION

5

In one embodiment, the present invention provides a TRIP1 nucleic acid molecule encoding a polypeptide selected from the group consisting of: the nucleic acid molecule of SEQ ID NO:1; the nucleic acid molecule of SEQ ID NO:2; a nucleic acid molecule encoding the polypeptide of SEQ ID NO:3, SEQ ID NO:4, or a biologically active fragment thereof; a nucleic acid molecule that encodes a polypeptide that is at least 70 percent identical to the polypeptide of SEQ ID NO:3 or SEQ ID NO:4; a nucleic acid molecule that hybridizes under stringent conditions to any of the above nucleic acids; and a nucleic acid molecule that is the complement of any of the above nucleic acids.

In another embodiment, the invention provides a nucleic acid molecule encoding amino acids 1-871 of the polypeptide of SEQ ID NO:3.

In one other embodiment, the invention provides vectors comprising the nucleic acids listed above, where the vectors can be amplification or expression vectors, suitable for use in prokaryotic or eukaryotic cells. Also provided are host cells comprising these vectors, wherein the host cells may be prokaryotic or eukaryotic cells.

The invention additionally provides a process for producing a TRIP1 polypeptide comprising the steps of expressing a polypeptide encoded by the nucleic acid of claim 1 in a suitable host and isolating the polypeptide, wherein the TRIP1 polypeptide may be SEQ ID NO:3, SEQ ID NO:4, or amino acids 1-871 of SEQ ID NO:3.

In yet another embodiment, the invention comprises a TRIP1 polypeptide selected from the group consisting of: the polypeptide of SEQ ID NO:3; the polypeptide that is amino acids 1-871 of SEQ ID NO:3; a  
5 polypeptide that is at least 70 percent identical to one of these polypeptides, or a polypeptide that is a biologically active fragment of one of these polypeptides.

In another embodiment, the present invention  
10 provides a TP2 nucleic acid molecule encoding a polypeptide selected from the group consisting of:

(a) the nucleic acid molecule of SEQ ID NO:13, SEQ ID NO:18 or SEQ ID NO:19;

(b) the nucleic acid molecule that is  
15 nucleotides 1920-2820 of SEQ ID NO:13;

(c) a nucleic acid molecule encoding the polypeptide of SEQ ID NO:14 or SEQ ID NO:20, or a biologically active fragment thereof;

(d) a nucleic acid molecule that encodes a  
20 polypeptide that is at least 90 percent identical to the polypeptide of SEQ ID NO:14 or SEQ ID NO:20;

(e) a nucleic acid molecule that hybridizes under stringent conditions to any of (a)-(d) above; and

(f) a nucleic acid molecule that is the  
25 complement of any of (a)-(e) above.

The invention further provides a nucleic acid molecule selected from the group consisting of:  
nucleotides 1-1689 of SEQ ID NO:13, nucleotides 1-1920  
30 of SEQ ID NO:13, nucleotides 1920-2820 of SEQ ID NO:13, nucleotides 2089-2820 of SEQ ID NO:13, and nucleotides 2089-2859 of SEQ ID NO:13.

Still further, the present invention provides a nucleic acid molecule encoding amino acids 640-940 of  
35 the polypeptide of SEQ ID NO:14 or the polypeptide of SEQ ID NO:20.

The invention also provides a process for producing a TP2 polypeptide comprising the steps of:

(a) expressing a polypeptide encoded by the nucleic acid of SEQ ID NO:13, SEQ ID NO:19 or a fragment thereof in a suitable host; and

(b) isolating the polypeptide, where the polypeptide may or may not possess an N-terminal methionine.

Yet further, the invention provides a TP2 polypeptide selected from the group consisting of: amino acids 1-563 of SEQ ID NO:14; amino acids 1-640 of SEQ ID NO:14; amino acids 640-940 of SEQ ID NO:14; amino acids 696-940 of SEQ ID NO:14; and amino acids 696-953 of SEQ ID NO:14.

Still further, the invention provides a method of increasing proliferation of a cell, comprising expressing a nucleic acid encoding TP2 or a biologically active fragment thereof, in the cell.

The invention also provides a method of increasing telomerase activity in a cell, comprising expressing a TP2 gene, or a biologically active fragment thereof, in the cell.

Additionally, the invention provides a method of decreasing telomerase in a cell, comprising expressing a TP2 mutant in a cell, wherein the mutant does not have TP2 biological activity.

Yet further, the invention provides a nucleic acid molecule encoding a mutant TP2 polypeptide, wherein the codon for aspartic acid at amino acid position 868 or 869 is changed to a codon for alanine, or wherein the codons for aspartic acid at amino acid positions 868 and 869 are changed to codons for alanine. The invention further provides polypeptides encoded by these nucleic acid molecules.

## BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A-1I depict the full length cDNA sequence of human TRIP1 (SEQ ID NO:1).

5

Figures 2A-2I depict the full length cDNA sequence of mouse TRIP1 (SEQ ID NO:2).

Figures 3A-3C depict the putative full length amino acid sequence (SEQ ID NO:3) of human TRIP1 as translated from the cDNA sequence.

10

Figures 4A-4C depict the putative full length amino acid sequence (SEQ ID NO:4) of mouse TRIP1 as translated from the cDNA sequence.

15

Figures 5A-5D depict the sequence of a cDNA encoding a large portion of human telomerase protein 2 ("TP2"; referred to in the Examples as "clone 32"; SEQ ID NO:13).

20

Figures 6A-6B depict the putative amino acid sequence of human TP2 (SEQ ID NO:14) as translated from the cDNA sequence of Figure 5.

25

Figure 7 (SEQ ID NO:18) depicts additional 3' sequence of human TP2 over the sequence set forth in Figure 5.

Figures 8A-8D (SEQ ID NO: 19) depict the full length human cDNA encoding TP2. This Figure combines the sequences of Figures 5 and 7.

30

Figures 9A-9B (SEQ ID NO: 20) depict the putative amino acid sequence of TP2 as translated from the cDNA sequence of Figure 8.

35

SECRET



Figure 10 is a schematic of the strategy used to make the TP2 mutant cDNA molecules. Nine PCR reactions were conducted to obtain the final TP2 cDNA mutant constructs. Six of these PCR reactions (indicated as 1-6) were primary reactions using the full length TP2 gene as a template. The final three PCR reactions (numbered 7-9) used the indicated PCR products from reactions 1-6 as templates. The oligonucleotide primers used for each PCR reaction are numbered according to their SEQ ID NOs.

Figures 11A-11C depict gels of telomerase assay results (Figs. 11A and 11B) and a Western blot (Fig. 11C). Abbreviations used in these Figures are set forth in Example 7A.

Figures 12A-12B depict a gel of telomerase assay results, and a Western blot, respectively. Abbreviations used in these Figures are set forth in Example 7B.

Figures 13A-13B depict a Western blot and a gel of telomerase assay results, respectively. Abbreviations used in these Figures are set forth in Example 7C.

Figure 14 depicts a gel of results of a telomerase assay from *in vitro* reconstitution of TP2 and telomerase RNA. Abbreviations used in this Figure are explained in Example 8.

Figure 15 depicts a gel of results of a telomerase assay from *in vitro* reconstitution of telomerase RNA plus wild type or mutant TP2. Abbreviations used in this Figure are explained in Example 8.

Figures 16A-16B depict a Western blot (16A) and a gel of a telomerase assay (16B) for cells transfected

with either wild type TP2 or a TP2 mutant. Details of the abbreviations used are described in Example 9.

Figures 17A-17B depict a gel of telomerase assay results (17A) and a Western blot (17B). Details of the abbreviations used are set forth in Example 8.

Figure 18 depicts a gel of telomerase assay results. Details of the abbreviations used are set forth in Example 8.

#### DETAILED DESCRIPTION OF THE INVENTION

Included in the scope of this invention are TRIP1 (referred to herein as "TRIP1") polypeptides such as the polypeptides of SEQ ID NO:3 and SEQ ID NO:4, and related biologically active polypeptide fragments and derivatives thereof. Also included within the scope of this invention are telomerase 2 (also referred to herein as "TP2") polypeptides such as the polypeptide of SEQ ID NO:14 and related biologically active polypeptide fragments and derivatives thereof. Further included within the scope of the present invention are nucleic acid molecules that encode these polypeptides, and methods for preparing the polypeptides. Such molecules may be useful as therapeutic agents in those cases where increasing TRIP1 activity or TP2 activity is desired.

In those situations in which TRIP1 and/or TP2 activity is to be decreased, such as in cancer cells in which TRIP1 activity and/or TP2 activity is elevated as compared to non-cancerous cells, TRIP1 and/or TP2 may serve as a target to identify a molecule which inhibits TRIP1 and/or TP2 activity, and/or a molecule which decreases or inhibits the protein-protein interaction of TRIP1 and TP2, or the binding of either TRIP1 or TP2 to telomerase RNA. Techniques that may be useful in

identifying such TRIP1 and/or TP2 inhibiting molecules are described in detail below. Alternatively, *ex vivo* or *in vivo* gene therapy may be employed to administer either TRIP1 or TP2 anti-sense molecules, or DNA  
5 constructs that may serve to disrupt or enhance TRIP1 and/or TP2 expression in the cells.

Also included within the scope of the present invention are non-human mammals such as mice, rats, rabbits, goats, or sheep in which the gene (or genes)  
10 encoding native TRIP1 and/or TP2 has been disrupted ("knocked out") such that the level of expression of this gene or genes is (are) significantly decreased or completely abolished. Such mammals may be prepared using techniques and methods such as those described in  
15 U.S. Patent No. 5,557,032. The present invention further includes non-human mammals such as mice, rats, rabbits, goats, or sheep in which the gene (or genes) encoding the TRIP1 and/or TP2 (either the native form of TRIP1 and/or TP2 for the mammal or a heterologous TRIP1  
20 and/or TP2 gene(s)) is (are) over expressed by the mammal, thereby creating a "transgenic" mammal. Such transgenic mammals may be prepared using well known methods such as those described in U.S. Patent No 5,489,743 and PCT patent application no. WO94/28122,  
25 published 8 December 1994. The present invention further includes non-human mammals in which either the TRIP1 or TP2 gene is knocked out, and the other gene (either TRIP1 or TP2) is over expressed.

The term "TRIP1 protein" or "TRIP1  
30 polypeptide" as used herein refers to any protein or polypeptide having the properties described herein for TRIP1. The small letter in front of the letters "TRIP1", when used, refers to a TRIP1 polypeptide from a particular mammal, *i.e.*, "hTRIP1" refers to human TRIP1,  
35 and "mTRIP1" refers to mouse TRIP1. The TRIP1

polypeptide may or may not have an amino terminal methionine, depending on the manner in which it is prepared. By way of illustration, TRIP1 protein or TRIP1 polypeptide refers to (1) an amino acid sequence  
5 encoded by TRIP1 nucleic acid molecules as defined in any of items (a)-(f) below, and biologically active peptide or polypeptide fragments derived therefrom, (2) naturally occurring allelic variants of the TRIP1 gene which result in one or more amino acid substitutions,  
10 deletions, and/or insertions as compared to the TRIP1 polypeptide of SEQ ID NO:3 or SEQ ID NO:4, and/or (3) chemically modified derivatives as well as nucleic acid and or amino acid sequence variants thereof as provided for herein.

15 As used herein, the term "TRIP1 fragment" refers to a peptide or polypeptide that is less than the full length amino acid sequence of naturally occurring TRIP1 protein but has substantially the same biological activity as TRIP1 polypeptide or TRIP1 protein described  
20 above. Such a fragment may be truncated at the amino terminus, the carboxy terminus, and/or internally, and may be chemically modified. Such TRIP1 fragments may be prepared with or without an amino terminal methionine.

As used herein, the term "TRIP1 derivative" or  
25 "TRIP1 variant" refers to a TRIP1 polypeptide, protein, or fragment that 1) has been chemically modified, as for example, by addition of one or more polyethylene glycol molecules, sugars, phosphates, or other such molecules not naturally attached to wild-type TRIP1 polypeptide,  
30 and/or 2) contains one or more nucleic acid or amino acid sequence substitutions, deletions, and/or insertions as compared to TRIP1 set forth in Figures 3 or 4.

As used herein, the terms "biologically TRIP1  
35 active polypeptide" and "biologically active TRIP1 fragment" refer to a TRIP1 peptide or polypeptide in

accordance with the above description for TRIP1 that has at least one of the following activities which have been identified for TRIP1: (1) specifically binding to telomerase RNA; and (2) binding to an antibody that is directed to an epitope on the polypeptide of SEQ ID NO:3 or SEQ ID NO:4.

As used herein, the term "TRIP1" when used to describe a nucleic acid molecule refers to a nucleic acid molecule or fragment thereof that (a) has the nucleotide sequence as set forth in SEQ ID NO:1 or SEQ ID NO:2; (b) has a nucleic acid sequence encoding a polypeptide that is at least 70 percent identical, but may be greater than 70 percent, *i.e.*, 80 percent, 90 percent, or even greater than 90 percent identical, to the polypeptide encoded by any of SEQ ID NOS:1 or 2; (c) is a naturally occurring allelic variant of (a) or (b); (d) is a nucleic acid variant of (a)-(c) produced as provided for herein; (e) has a sequence that is complementary to (a)-(d); and/or (f) hybridizes to any of (a)-(e) under stringent conditions.

The terms "telomerase protein 2", "TP2 protein" or "TP2 polypeptide" as used herein refer to any protein or polypeptide having the properties described herein for TP2. The small letter in front of the letters "TP2", when used, refers to a TP2 polypeptide from a particular mammal, *i.e.*, "hTP2" refers to human TP2, and "mTP2" refers to mouse TP2. The TP2 polypeptide may or may not have an amino terminal methionine, depending on the manner in which it is prepared. By way of illustration, TP2 protein or T2 polypeptide refers to (1) an amino acid sequence encoded by TP2 nucleic acid molecules as defined in SEQ ID NO:13, SEQ ID NO:18, and SEQ ID NO:19 herein, and biologically active peptide or polypeptide fragments derived therefrom such as, for example, a peptide

encoded by nucleotides 1950-2888 of SEQ ID NO:13, (2) naturally occurring allelic variants of the TP2 gene which result in one or more amino acid substitutions, deletions, and/or insertions as compared to the TP2 polypeptides of SEQ ID NO:14 or SEQ ID NO:20, and/or (3) chemically modified derivatives as well as nucleic acid and or amino acid sequence variants thereof as provided for herein.

As used herein, the term "TP2 fragment" refers to a peptide or polypeptide that is less than the full length amino acid sequence of naturally occurring TP2 protein but has substantially the same biological activity as TP2 polypeptide or TP2 protein described above. Such a fragment may be truncated at the amino terminus, the carboxy terminus, and/or internally, and may be chemically modified. Such TP2 fragments may be prepared with or without an amino terminal methionine.

As used herein, the term "TP2 derivative" or "TP2 variant" refers to a TP2 polypeptide, protein, or fragment that 1) has been chemically modified, as for example, by addition of one or more polyethylene glycol molecules, sugars, phosphates, or other such molecules not naturally attached to wild-type TP2 polypeptide, and/or 2) contains one or more nucleic acid or amino acid sequence substitutions, deletions, and/or insertions as compared to the TP2 amino acid sequences set forth in Figures 6 and 9. Preferred TP2 fragments include amino acids 1-563 of SEQ ID NO:14; amino acids 1-640 of SEQ ID NO:14; amino acids 640-940 of SEQ ID NO:14; amino acids 696-940 of SEQ ID NO:14; and amino acids 696-953 of SEQ ID NO:14.

As used herein, the terms "biologically active TP2 polypeptide" and "biologically active TP2 fragment" refer to a TP2 peptide or polypeptide in accordance with the above description for TP2, where the TP2 also has

5 catalytic activity in a telomerase assay. In addition,  
the biologically active TP2 has at least one of the  
following properties which have been identified for TP2:  
binds to an antibody that is directed to an epitope on  
the TP2 polypeptide of SEQ ID NO:14 or SEQ ID NO:20, and  
also either (a) specifically interacts with other  
telomerase protein components and/or the RNA component  
of the telomerase complex; (b) contains one or more  
identifiable reverse transcriptase motifs in its amino  
10 acid sequence; or (c) possesses the properties set forth  
in both (a) and (b).

As used herein, the term "TP2" when used to  
describe a nucleic acid molecule refers to a nucleic  
acid molecule or fragment thereof that (a) has the  
15 nucleotide sequence that as set forth in SEQ ID NO:13 or  
SEQ ID NO: 19, or a fragment thereof that is less than  
the full length SEQ ID NO:13 or SEQ ID NO:19; (b) has a  
nucleic acid sequence encoding a polypeptide that is at  
least 70 percent identical, but may be greater than 70  
20 percent, *i.e.*, 80 percent, 90 percent, or even greater  
than 90 percent identical, to the polypeptide encoded by  
SEQ ID 13, or the polypeptide encoded by SEQ ID NO: 19;  
(c) is a naturally occurring allelic variant of (a) or  
(b); (d) is a nucleic acid variant of (a)-(c) produced  
25 as provided for herein; (e) has a sequence that is  
complementary to (a)-(d); and/or (f) hybridizes to any  
of (a)-(e) under stringent conditions. Preferred TP2  
nucleic acids of the present invention include full  
length TP2 as set forth in SEQ ID NO:13, 1-1689 of SEQ  
30 ID NO:13, nucleotides 1-1920 of SEQ ID NO:13,  
nucleotides 1920-2820 of SEQ ID NO:13, nucleotides 2089-  
2820 of SEQ ID NO:13, and nucleotides 2089-2859 of SEQ  
ID NO:13.

Percent sequence identity can be determined by standard methods that are commonly used to compare the similarity in position of the amino acids of two polypeptides. By way of example, using a computer program such as BLAST or FASTA, the two polypeptides for which the percent sequence identity is to be determined are aligned for optimal matching of their respective amino acids (the "matched span", which can include the full length of one or both sequences, or a pre-determined portion of one or both sequences). Each computer program provides a "default" opening penalty and a "default" gap penalty, and a scoring matrix such as PAM 250. A standard scoring matrix (see Dayhoff et al., in: *Atlas of Protein Sequence and Structure*, vol. 5, supp.3 [1978]) can be used in conjunction with the computer program. The percent identity can then be calculated by determining the percent identity using an algorithm contained in a program such as FASTA:

$$\frac{\text{Total number of identical matches}}{[\text{length of the longer sequence within the matched span}] + [\text{number of gaps introduced into the longer sequence in order to align the two sequences}]} \times 100$$

Polypeptides that are at least 70 percent identical will typically have one or more amino acid substitutions, deletions, and/or insertions as compared with wild type TRIP1. Usually, the substitutions will be conservative so as to have little or no effect on the overall net charge, polarity, or hydrophobicity of the protein but optionally may increase the activity of TRIP1. Conservative substitutions are set forth in Table I below.



Table I  
Conservative Amino Acid Substitutions

Basic:	arginine
	lysine
	histidine
Acidic:	glutamic acid
	aspartic acid
Polar:	glutamine
	asparagine
Hydrophobic:	leucine
	isoleucine
	valine
Aromatic:	phenylalanine
	tryptophan
	tyrosine
Small:	glycine
	alanine
	serine
	threonine
	methionine

5           The term "stringent conditions" refers to  
hybridization and washing under conditions that permit  
only binding of a nucleic acid molecule such as an  
oligonucleotide or cDNA molecule probe to highly  
homologous sequences. One stringent wash solution is  
10 0.015 M NaCl, 0.005 M NaCitrate, and 0.1 percent SDS  
used at a temperature of 55°C-65°C. Another stringent  
wash solution is 0.2 X SSC and 0.1 percent SDS used at a  
temperature of between 50°C-65°C. Where oligonucleotide  
probes are used to screen cDNA or genomic libraries, the  
15 following stringent washing conditions may be used. One  
protocol uses 6 X SSC with 0.05 percent sodium  
pyrophosphate at a temperature of 35°C-62°C, depending  
on the length of the oligonucleotide probe. For

example, 14 base pair probes are washed at 35-40°C, 17  
base pair probes are washed at 45-50°C, 20 base pair  
probes are washed at 52-57°C, and 23 base pair probes  
are washed at 57-63°C. The temperature can be increased  
5 2-3°C where the background non-specific binding appears  
high. A second protocol utilizes tetramethylammonium  
chloride (TMAC) for washing oligonucleotide probes. One  
stringent washing solution is 3 M TMAC, 50 mM Tris-HCl,  
pH 8.0, and 0.2 percent SDS. The washing temperature  
10 using this solution is a function of the length of the  
probe. For example, a 17 base pair probe is washed at  
about 45-50°C.

As used herein, the terms "effective amount"  
and "therapeutically effective amount" refer to the  
15 amount of TRIP1 and/or TP2 necessary to support one or  
more biological activities of TRIP1 and/or TP2 as set  
forth above.

The TRIP1 and/or TP2 polypeptides that have  
use in practicing the present invention may be naturally  
20 occurring full length polypeptides, or truncated  
polypeptides or peptides (*i.e.*, "fragments"). The  
polypeptides or fragments may be chemically modified,  
*i.e.*, glycosylated, phosphorylated, and/or linked to a  
polymer, as described below, and they may have an amino  
25 terminal methionine, depending on how they are prepared.  
In addition, the polypeptides or fragments may be  
variants of the naturally occurring TRIP1 and/or TP2  
polypeptide (*i.e.*, may contain one or more amino acid  
deletions, insertions, and/or substitutions as compared  
30 with naturally occurring TRIP1 or TP2).

The full length TRIP1 or TP2 polypeptide or  
fragment thereof can be prepared using well known  
recombinant DNA technology methods such as those set  
forth in Sambrook *et al.* (*Molecular Cloning: A*  
35 *Laboratory Manual*, Cold Spring Harbor Laboratory Press,  
Cold Spring Harbor, NY [1989]) and/or Ausubel *et al.*,

eds, (*Current Protocols in Molecular Biology*, Green Publishers Inc. and Wiley and Sons, NY [1994]). A gene or cDNA encoding the TRIP1 or TP2 protein or fragment thereof may be obtained for example by screening a  
5 genomic or cDNA library, or by PCR amplification. Alternatively, a gene encoding the TRIP1 or TP2 polypeptide or fragment may be prepared by chemical synthesis using methods well known to the skilled artisan such as those described by Engels *et al.* (*Angew. Chem. Intl. Ed.*, 28:716-734 [1989]). These methods  
10 include, *inter alia*, the phosphotriester, phosphoramidite, and H-phosphonate methods for nucleic acid synthesis. A preferred method for such chemical synthesis is polymer-supported synthesis using standard  
15 phosphoramidite chemistry. Typically, the DNA encoding the TRIP1 or TP2 polypeptide will be several hundred nucleotides in length. Nucleic acids larger than about 100 nucleotides can be synthesized as several fragments using these methods. The fragments can then be ligated  
20 together to form the full length TRIP1 or TP2 polypeptide. Usually, the DNA fragment encoding the amino terminus of the polypeptide will have an ATG, which encodes a methionine residue. This methionine may or may not be present on the mature form of the TRIP1 or  
25 TP2 polypeptide, depending on whether the polypeptide produced in the host cell is secreted from that cell.

In some cases, it may be desirable to prepare nucleic acid and/or amino acid variants of naturally occurring TRIP1 or TP2. Nucleic acid variants (wherein  
30 one or more nucleotides are designed to differ from the wild-type or naturally occurring TRIP1 or TP2) may be produced using site directed mutagenesis or PCR amplification where the primer(s) have the desired point mutations (see Sambrook *et al.*, *supra*, and Ausubel *et al.*, *supra*, for descriptions of mutagenesis techniques).  
35 Chemical synthesis using methods described by Engels *et*

al., *supra*, may also be used to prepare such variants. Other methods known to the skilled artisan may be used as well. Preferred nucleic acid variants are those containing nucleotide substitutions accounting for codon  
5 preference in the host cell that is to be used to produce TRIP1 or TP2. Other preferred variants are those encoding conservative amino acid changes as described above (e.g., wherein the charge or polarity of the naturally occurring amino acid side chain is not  
10 altered substantially by substitution with a different amino acid) as compared to wild type, and/or those designed to either generate a novel glycosylation and/or phosphorylation site(s) on TRIP1 or TP2, or those designed to delete an existing glycosylation and/or  
15 phosphorylation site(s) on TRIP1 or TP2.

The TRIP1 or TP2 gene or cDNA can be inserted into an appropriate expression vector for expression in a host cell. The vector is typically selected to be functional in the particular host cell employed (*i.e.*,  
20 the vector is compatible with the host cell machinery such that amplification of the TRIP1 or TP2 gene and/or expression of the gene can occur). The TRIP1 or TP2 polypeptide or fragment thereof may be amplified/expressed in prokaryotic, yeast, insect  
25 (baculovirus systems) and/or eukaryotic host cells. Selection of the host cell will depend at least in part on whether the TRIP1 or TP2 polypeptide or fragment thereof is to be glycosylated and/or phosphorylated. If so, yeast, insect, or mammalian host cells are  
30 preferable; yeast cells can typically glycosylate and phosphorylate the polypeptide, and insect and mammalian cells can glycosylate and/or phosphorylate the polypeptide as it naturally occurs on the TRIP1 or TP2 polypeptide (*i.e.*, "native" glycosylation and/or  
35 phosphorylation).

Typically, the vectors used in any of the host cells will contain 5' flanking sequence (also referred to as a "promoter") and other regulatory elements as well such as an enhancer(s), an origin of replication element, a transcriptional termination element, a complete intron sequence containing a donor and acceptor splice site, a signal peptide sequence, a ribosome binding site element, a polyadenylation sequence, a polylinker region for inserting the nucleic acid encoding the polypeptide to be expressed, and a selectable marker element. Each of these elements is discussed below. Optionally, the vector may contain a "tag" sequence, *i.e.*, an oligonucleotide sequence located at the 5' or 3' end of the TRIP1 or TP2 coding sequence that encodes polyHis (such as hexaHis) or another small immunogenic sequence. This tag will be expressed along with the protein, and can serve as an affinity tag for purification of the TRIP1 or TP2 polypeptide from the host cell. Optionally, the tag can subsequently be removed from the purified TRIP1 or TP2 polypeptide by various means such as using a selected peptidase for example.

The 5' flanking sequence may be homologous (*i.e.*, from the same species and/or strain as the host cell), heterologous (*i.e.*, from a species other than the host cell species or strain), hybrid (*i.e.*, a combination of 5' flanking sequences from more than one source), synthetic, or it may be the native TRIP1 or TP2 5' flanking sequence. As such, the source of the 5' flanking sequence may be any unicellular prokaryotic or eukaryotic organism, any vertebrate or invertebrate organism, or any plant, provided that the 5' flanking sequence is functional in, and can be activated by, the host cell machinery.

The 5' flanking sequences useful in the vectors of this invention may be obtained by any of

several methods well known in the art. Typically, 5' flanking sequences useful herein other than the TRIP1 or TP2 5' flanking sequence will have been previously identified by mapping and/or by restriction endonuclease digestion and can thus be isolated from the proper tissue source using the appropriate restriction endonucleases. In some cases, the full nucleotide sequence of the 5' flanking sequence may be known. Here, the 5' flanking sequence may be synthesized using the methods described above for nucleic acid synthesis or cloning.

Where all or only a portion of the 5' flanking sequence is known, it may be obtained using PCR and/or by screening a genomic library with suitable oligonucleotide and/or 5' flanking sequence fragments from the same or another species.

Where the 5' flanking sequence is not known, a fragment of DNA containing a 5' flanking sequence may be isolated from a larger piece of DNA that may contain, for example, a coding sequence or even another gene or genes. Isolation may be accomplished by restriction endonuclease digestion using one or more carefully selected enzymes to isolate the proper DNA fragment. After digestion, the desired fragment may be isolated by agarose gel purification, Qiagen® column or other methods known to the skilled artisan. Selection of suitable enzymes to accomplish this purpose will be readily apparent to one of ordinary skill in the art.

The origin of replication element is typically a part of prokaryotic expression vectors purchased commercially, and aids in the amplification of the vector in a host cell. Amplification of the vector to a certain copy number can, in some cases, be important for optimal expression of the TRIP1 or TP2 polypeptide. If the vector of choice does not contain an origin of

replication site, one may be chemically synthesized based on a known sequence, and ligated into the vector.

The transcription termination element is typically located 3' of the end of the TRIP1 or TP2 polypeptide coding sequence and serves to terminate transcription of the TRIP1 or TP2 polypeptide. Usually, the transcription termination element in prokaryotic cells is a G-C rich fragment followed by a poly T sequence. While the element is easily cloned from a library or even purchased commercially as part of a vector, it can also be readily synthesized using methods for nucleic acid synthesis such as those described above.

A selectable marker gene element encodes a protein necessary for the survival and growth of a host cell grown in a selective culture medium. Typical selection marker genes encode proteins that (a) confer resistance to antibiotics or other toxins, e.g., ampicillin, tetracycline, or kanamycin for prokaryotic host cells, (b) complement auxotrophic deficiencies of the cell; or (c) supply critical nutrients not available from complex media. Preferred selectable markers are the kanamycin resistance gene, the ampicillin resistance gene, and the tetracycline resistance gene.

The ribosome binding element, commonly called the Shine-Dalgarno sequence (prokaryotes) or the Kozak sequence (eukaryotes), is necessary for translation initiation of mRNA. The element is typically located 3' to the promoter and 5' to the coding sequence of the TRIP1 or TP2 polypeptide to be synthesized. The Shine-Dalgarno sequence is varied but is typically a polypurine (*i.e.*, having a high A-G content). Many Shine-Dalgarno sequences have been identified, each of which can be readily synthesized using methods set forth above and used in a prokaryotic vector.

08951733 101697  
SECRET

In those cases where it is desirable for TRIP1 or TP2 to be secreted from the host cell, a signal sequence may be used to direct the TRIP1 or TP2 polypeptide out of the host cell where it is synthesized, and the carboxy-terminal part of the protein may be deleted in order to prevent membrane anchoring. Typically, the signal sequence is positioned in the coding region of TRIP1 or TP2 nucleic acid sequence, or directly at the 5' end of the TRIP1 or TP2 coding region. Many signal sequences have been identified, and any of them that are functional in the selected host cell may be used in conjunction with the TRIP1 or TP2 gene. Therefore, the signal sequence may be homologous or heterologous to the TRIP1 or TP2 polypeptide, and may be homologous or heterologous to the TRIP1 or TP2 polypeptide. Additionally, the signal sequence may be chemically synthesized using methods set forth above. In most cases, secretion of the polypeptide from the host cell via the presence of a signal peptide will result in the removal of the amino terminal methionine from the polypeptide.

In many cases, transcription of the TRIP1 or TP2 polypeptide is increased by the presence of one or more introns on the vector; this is particularly true where TRIP1 or TP2 is produced in eukaryotic host cells, especially mammalian host cells. The introns used may be naturally occurring within the TRIP1 or TP2 nucleic acid sequence, especially where the TRIP1 or TP2 sequence used is a full length genomic sequence or a fragment thereof. Where the intron is not naturally occurring within the TRIP1 or TP2 DNA sequence (as for most cDNAs), the intron(s) may be obtained from another source. The position of the intron with respect to the 5' flanking sequence and the TRIP1 or TP2 coding sequence is important, as the intron must be transcribed to be effective. As such, where the TRIP1 or TP2



nucleic acid sequence is a cDNA sequence, the preferred position for the intron is 3' to the transcription start site, and 5' to the polyA transcription termination sequence. Preferably for TRIP1 or TP2 cDNAs, the intron will be located on one side or the other (*i.e.*, 5' or 3') of the TRIP1 or TP2 coding sequence such that it does not interrupt the this coding sequence. Any intron from any source, including any viral, prokaryotic and eukaryotic (plant or animal) organisms, may be used to practice this invention, provided that it is compatible with the host cell(s) into which it is inserted. Also included herein are synthetic introns. Optionally, more than one intron may be used in the vector.

Where one or more of the elements set forth above are not already present in the vector to be used, they may be individually obtained and ligated into the vector. Methods used for obtaining each of the elements are well known to the skilled artisan and are comparable to the methods set forth above (*i.e.*, synthesis of the DNA, library screening, and the like).

The final vectors used to practice this invention are typically constructed from a starting vectors such as a commercially available vector. Such vectors may or may not contain some of the elements to be included in the completed vector. If none of the desired elements are present in the starting vector, each element may be individually ligated into the vector by cutting the vector with the appropriate restriction endonuclease(s) such that the ends of the element to be ligated in and the ends of the vector are compatible for ligation. In some cases, it may be necessary to "blunt" the ends to be ligated together in order to obtain a satisfactory ligation. Blunting is accomplished by first filling in "sticky ends" using Klenow DNA polymerase or T4 DNA polymerase in the presence of all four nucleotides. This procedure is well known in the

art and is described for example in Sambrook *et al.*,  
*supra*.

Alternatively, two or more of the elements to  
be inserted into the vector may first be ligated  
5 together (if they are to be positioned adjacent to each  
other) and then ligated into the vector.

One other method for constructing the vector  
to conduct all ligations of the various elements  
simultaneously in one reaction mixture. Here, many  
10 nonsense or nonfunctional vectors will be generated due  
to improper ligation or insertion of the elements,  
however the functional vector may be identified and  
selected by restriction endonuclease digestion.

Preferred vectors for practicing this  
15 invention are those which are compatible with bacterial,  
insect, and mammalian host cells. Such vectors include,  
*inter alia*, pCRII and pCR3 (Invitrogen Company, San  
Diego, CA), pBSII (Stratagene Company, LaJolla, CA), and  
pETL (BlueBacII; Invitrogen).

20 After the vector has been constructed and a  
TRIP1 nucleic acid has been inserted into the proper  
site of the vector, the completed vector may be inserted  
into a suitable host cell for amplification and/or TRIP1  
or TP2 polypeptide expression.

25 Host cells may be prokaryotic host cells (such  
as *E. coli*) or eukaryotic host cells (such as a yeast  
cell, an insect cell, or a vertebrate cell). The host  
cell, when cultured under appropriate conditions, can  
synthesize TRIP1 or TP2 protein which can subsequently  
30 be collected from the culture medium (if the host cell  
secretes it into the medium) or directly from the host  
cell producing it (if it is not secreted). After  
collection, the TRIP1 or TP2 protein can be purified  
using methods such as molecular sieve chromatography,  
35 affinity chromatography, and the like.

Selection of the host cell will depend in part on whether the TRIP1 or TP2 protein is to be glycosylated or phosphorylated (in which case eukaryotic host cells are preferred), and the manner in which the host cell is able to "fold" the protein into its native tertiary structure (e.g., proper orientation of disulfide bridges, etc.) such that biologically active protein is prepared by the cell. However, where the host cell does not synthesize biologically active TRIP1 or TP2, the TRIP1 or TP2 may be "folded" after synthesis using appropriate chemical conditions as discussed below.

Suitable cells or cell lines may be mammalian cells, such as Chinese hamster ovary cells (CHO) or 3T3 cells. The selection of suitable mammalian host cells and methods for transformation, culture, amplification, screening and product production and purification are known in the art. Other suitable mammalian cell lines, are the monkey COS-1 and COS-7 cell lines, and the CV-1 cell line. Further exemplary mammalian host cells include primate cell lines and rodent cell lines, including transformed cell lines. Normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, as well as primary explants, are also suitable. Candidate cells may be genotypically deficient in the selection gene, or may contain a dominantly acting selection gene. Other suitable mammalian cell lines include but are not limited to, mouse neuroblastoma N2A cells, HeLa, mouse L-929 cells, 3T3 lines derived from Swiss, Balb-c or NIH mice, BHK or HaK hamster cell lines.

Similarly useful as host cells suitable for the present invention are bacterial cells. For example, the various strains of *E. coli* (e.g., HB101, DH5 $\alpha$ , DH10, and MC1061) are well-known as host cells in the field of biotechnology. Various strains of *B. subtilis*,

*Pseudomonas spp.*, other *Bacillus spp.*, *Streptomyces spp.*, and the like may also be employed in this method.

Many strains of yeast cells known to those skilled in the art are also available as host cells for expression of the polypeptides of the present invention. Additionally, where desired, insect cells may be utilized as host cells in the method of the present invention (Miller *et al.*, *Genetic Engineering* 8:277-298 [1986]).

Insertion (also referred to as "transformation" or "transfection") of the vector into the selected host cell may be accomplished using such methods as calcium chloride, electroporation, microinjection, lipofection or the DEAE-dextran method. The method selected will in part be a function of the type of host cell to be used. These methods and other suitable methods are well known to the skilled artisan, and are set forth, for example, in Sambrook *et al.*, *supra*.

The host cells containing the vector (*i.e.*, transformed or transfected) may be cultured using standard media well known to the skilled artisan. The media will usually contain all nutrients necessary for the growth and survival of the cells. Suitable media for culturing *E. coli* cells are for example, Luria Broth (LB) and/or Terrific Broth (TB). Suitable media for culturing eukaryotic cells are RPMI 1640, MEM, DMEM, all of which may be supplemented with serum and/or growth factors as required by the particular cell line being cultured. A suitable medium for insect cultures is Grace's medium supplemented with yeastolate, lactalbumin hydrolysate, and/or fetal calf serum as necessary.

Typically, an antibiotic or other compound useful for selective growth of the transformed cells only is added as a supplement to the media. The compound to be used will be dictated by the selectable

marker element present on the plasmid with which the host cell was transformed. For example, where the selectable marker element is kanamycin resistance, the compound added to the culture medium will be kanamycin.

5           The amount of TRIP1 or TP2 polypeptide produced in the host cell can be evaluated using standard methods known in the art. Such methods include, without limitation, Western blot analysis, SDS-polyacrylamide gel electrophoresis, non-denaturing gel  
10 electrophoresis, HPLC separation, immunoprecipitation, and/or activity assays such as DNA binding gel shift assays.

          If the TRIP1 or TP2 polypeptide has been designed to be secreted from the host cells, the  
15 majority of polypeptide may be found in the cell culture medium. Polypeptides prepared in this way will typically not possess an amino terminal methionine, as it is removed during secretion from the cell. If  
20 however, the TRIP1 or TP2 polypeptide is not secreted from the host cells, it will be present in the cytoplasm (for eukaryotic, gram positive bacteria, and insect host cells) or in the periplasm (for gram negative bacteria host cells) and may have an amino terminal methionine.

          For intracellular TRIP1 or TP2 protein, the  
25 host cells are typically first disrupted mechanically or osmotically to release the cytoplasmic contents into a buffered solution. TRIP1 or TP2 polypeptide can then be isolated from this solution.

          Purification of TRIP1 polypeptide from  
30 solution can be accomplished using a variety of techniques. If the polypeptide has been synthesized such that it contains a tag such as Hexahistidine (TRIP1/hexaHis or TP2/hexaHis) or other small peptide at  
35 either its carboxyl or amino terminus, it may essentially be purified in a one-step process by passing the solution through an affinity column where the column

089473-10167  
SECRET

matrix has a high affinity for the tag or for the polypeptide directly (i.e., a monoclonal antibody specifically recognizing TRIP1 or TP2). For example, polyhistidine binds with great affinity and specificity to nickel, thus an affinity column of nickel (such as the Qiagen nickel columns) can be used for purification of TRIP1/polyHis or TP2/polyHis. (See for example, Ausubel et al., eds., *Current Protocols in Molecular Biology*, Section 10.11.8, John Wiley & Sons, New York [1993]).

Where the TRIP1 or TP2 polypeptide has no tag and no antibodies are available, other well known procedures for purification can be used. Such procedures include, without limitation, ion exchange chromatography, molecular sieve chromatography, HPLC, native gel electrophoresis in combination with gel elution, and preparative isoelectric focusing ("Isoprime" machine/technique, Hoefer Scientific). In some cases, two or more of these techniques may be combined to achieve increased purity.

If it is anticipated that the TRIP1 or TP2 polypeptide will be found primarily in the periplasmic space of the bacteria or the cytoplasm of eukaryotic cells, the contents of the periplasm or cytoplasm, including inclusion bodies (e.g., gram-negative bacteria) if the processed polypeptide has formed such complexes, can be extracted from the host cell using any standard technique known to the skilled artisan. For example, the host cells can be lysed to release the contents of the periplasm by French press, homogenization, and/or sonication. The homogenate can then be centrifuged.

If the TRIP1 or TP2 polypeptide has formed inclusion bodies in the periplasm, the inclusion bodies can often bind to the inner and/or outer cellular membranes and thus will be found primarily in the pellet

material after centrifugation. The pellet material can then be treated with a chaotropic agent such as guanidine or urea to release, break apart, and solubilize the inclusion bodies. The TRIP1 or TP2 polypeptide in its now soluble form can then be analyzed using gel electrophoresis, immunoprecipitation or the like. If it is desired to isolate the TRIP1 or TP2 polypeptide, isolation may be accomplished using standard methods such as those set forth below and in Marston et al. (*Meth. Enz.*, 182:264-275 [1990]).

If TRIP1 or TP2 polypeptide inclusion bodies are not formed to a significant degree in the periplasm of the host cell, the TRIP1 or TP2 polypeptide will be found primarily in the supernatant after centrifugation of the cell homogenate, and the TRIP1 or TP2 polypeptide can be isolated from the supernatant using methods such as those set forth below.

In those situations where it is preferable to partially or completely isolate the TRIP1 or TP2 polypeptide, purification can be accomplished using standard methods well known to the skilled artisan. Such methods include, without limitation, separation by electrophoresis followed by electroelution, various types of chromatography (immunoaffinity, molecular sieve, and/or ion exchange), and/or high pressure liquid chromatography. In some cases, it may be preferable to use more than one of these methods for complete purification.

In addition to preparing and purifying TRIP1 or TP2 polypeptide using recombinant DNA techniques, the TRIP1 or TP2 polypeptides, fragments, and/or derivatives thereof may be prepared by chemical synthesis methods (such as solid phase peptide synthesis) using methods known in the art such as those set forth by Merrifield et al., (*J. Am. Chem. Soc.*, 85:2149 [1964]), Houghten et al. (*Proc Natl Acad. Sci. USA*, 82:5132 [1985]), and

Stewart and Young (Solid Phase Peptide Synthesis, Pierce Chem Co, Rockford, IL [1984]). Such polypeptides may be synthesized with or without a methionine on the amino terminus. Chemically synthesized TRIP1 or TP2

5 polypeptides or fragments may be oxidized using methods set forth in these references to form disulfide bridges. The TRIP1 or TP2 polypeptides or fragments may be employed as biologically active or immunological substitutes for natural, purified TRIP1 or TP2  
10 polypeptides in therapeutic and immunological processes.

Chemically modified TRIP1 or TP2 compositions (*i.e.*, "derivatives") where the TRIP1 or TP2 polypeptide is linked to a polymer ("TRIP1 or TP2 -polymers") are included within the scope of the present invention. The  
15 polymer selected is typically water soluble so that the protein to which it is attached does not precipitate in an aqueous environment, such as a physiological environment. The polymer selected is usually modified to have a single reactive group, such as an active ester  
20 for acylation or an aldehyde for alkylation, so that the degree of polymerization may be controlled as provided for in the present methods. A preferred reactive aldehyde is polyethylene glycol propionaldehyde, which is water stable, or mono C1-C10 alkoxy or aryloxy  
25 derivatives thereof (see U.S. Patent 5,252,714). The polymer may be branched or unbranched. Included within the scope of TRIP1-polymers or TP2-polymers is a mixture of polymers. Preferably, for therapeutic use of the end-product preparation, the polymer will be  
30 pharmaceutically acceptable. The water soluble polymer or mixture thereof may be selected from the group consisting of, for example, polyethylene glycol (PEG), monomethoxy-polyethylene glycol, dextran, cellulose, or other carbohydrate based polymers, poly-(N-vinyl  
35 pyrrolidone) polyethylene glycol, propylene glycol homopolymers, a polypropylene oxide/ethylene oxide co-



polymer, polyoxyethylated polyols (e.g., glycerol) and polyvinyl alcohol. For the acylation reactions, the polymer(s) selected should have a single reactive ester group. For reductive alkylation, the polymer(s) selected should have a single reactive aldehyde group. The polymer may be of any molecular weight, and may be branched or unbranched.

Pegylation of TRIP1 or TP2 may be carried out by any of the pegylation reactions known in the art, as described for example in the following references: *Focus on Growth Factors* 3: 4-10 (1992); EP 0 154 316; and EP 0 401 384. Preferably, the pegylation is carried out via an acylation reaction or an alkylation reaction with a reactive polyethylene glycol molecule (or an analogous reactive water-soluble polymer) as described below.

Pegylation by acylation generally involves reacting an active ester derivative of polyethylene glycol (PEG) with an TRIP1 or TP2 protein. Any known or subsequently discovered reactive PEG molecule may be used to carry out the pegylation of TRIP1 or TP2. A preferred activated PEG ester is PEG esterified to N-hydroxysuccinimide ("NHS"). As used herein, "acylation" is contemplated to include without limitation the following types of linkages between TRIP1 or TP2 and a water soluble polymer such as PEG: amide, carbamate, urethane, and the like, as described in *Bioconjugate Chem.* 5:133-140 (1994). Reaction conditions may be selected from any of those known in the pegylation art or those subsequently developed, provided that conditions such as temperature, solvent, and pH that would inactivate the TRIP1 or TP2 species to be modified are avoided.

Pegylation by acylation usually results in a poly-pegylated TRIP1 or TP2 product, wherein the lysine  $\epsilon$ -amino groups are pegylated via an acyl linking group. Preferably, the connecting linkage will be an amide.

Also preferably, the resulting product will be at least about 95 percent mono, di- or tri- pegylated. However, some species with higher degrees of pegylation (up to the maximum number of lysine  $\epsilon$ -amino acid groups of TRIP1 or TP2 plus one  $\alpha$ -amino group at the amino terminus of TRIP1 or TP2) will normally be formed in amounts depending on the specific reaction conditions used. If desired, more purified pegylated species may be separated from the mixture, particularly unreacted species, by standard purification techniques, including, among others, dialysis, salting-out, ultrafiltration, ion-exchange chromatography, gel filtration chromatography and electrophoresis.

Pegylation by alkylation generally involves reacting a terminal aldehyde derivative of PEG with a protein such as TRIP1 or TP2 in the presence of a reducing agent. Regardless of the degree of pegylation, the PEG groups are preferably attached to the protein via a  $-\text{CH}_2\text{-NH}-$  group. With particular reference to the  $-\text{CH}_2-$  group, this type of linkage is referred to herein as an "alkyl" linkage.

Derivatization via reductive alkylation to produce a monopegylated product exploits the differential reactivity of different types of primary amino groups (lysine versus the N-terminal) available for derivatization in TRIP1 or TP2. Typically, the reaction is performed at a pH (see below) which allows one to take advantage of the  $\text{pK}_a$  differences between the  $\epsilon$ -amino groups of the lysine residues and that of the  $\alpha$ -amino group of the N-terminal residue of the protein. By such selective derivatization, attachment of a water soluble polymer that contains a reactive group such as an aldehyde, to a protein is controlled: the conjugation with the polymer occurs predominantly at the N-terminus of the protein without significant modification of other reactive groups such as the lysine side chain amino

groups. The present invention provides for a substantially homogeneous preparation of TRIP1-monopolymer or TP2-monopolymer protein conjugate molecules (meaning TRIP1 or TP2 protein to which a polymer molecule has been attached substantially only (*i.e.*, at least about 95%) in a single location on the TRIP1 or TP2 protein. More specifically, if polyethylene glycol is used, the present invention also provides for pegylated TRIP1 or TP2 protein lacking possibly antigenic linking groups, and having the polyethylene glycol molecule directly coupled to the TRIP1 or TP2 protein.

A particularly preferred water-soluble polymer for use herein is polyethylene glycol, abbreviated PEG. As used herein, polyethylene glycol is meant to encompass any of the forms of PEG that have been used to derivatize other proteins, such as mono-(C1-C10) alkoxy- or aryloxy-polyethylene glycol.

In general, chemical derivatization may be performed under any suitable conditions used to react a biologically active substance with an activated polymer molecule. Methods for preparing pegylated TRIP1 or TP2 will generally comprise the steps of (a) reacting an TRIP1 or TP2 polypeptide with polyethylene glycol (such as a reactive ester or aldehyde derivative of PEG) under conditions whereby TRIP1 or TP2 becomes attached to one or more PEG groups, and (b) obtaining the reaction product(s). In general, the optimal reaction conditions for the acylation reactions will be determined based on known parameters and the desired result. For example, the larger the ratio of PEG: protein, the greater the percentage of poly-pegylated product.

Reductive alkylation to produce a substantially homogeneous population of mono-polymer/TRIP1 or TP2 protein conjugate molecule will generally comprise the steps of: (a) reacting a TRIP1 or

TP2 protein with a reactive PEG molecule under reductive alkylation conditions, at a pH suitable to permit selective modification of the  $\alpha$ -amino group at the amino terminus of the TRIP1 or TP2 protein; and (b) obtaining  
5 the reaction product(s).

For a substantially homogeneous population of mono-polymer/TRIP1 or TP2 protein conjugate molecules, the reductive alkylation reaction conditions are those which permit the selective attachment of the water  
10 soluble polymer moiety to the N-terminus of TRIP1 or TP2. Such reaction conditions generally provide for  $pK_a$  differences between the lysine amino groups and the  $\alpha$ -amino group at the N-terminus (the  $pK_a$  being the pH at which 50% of the amino groups are protonated and 50% are  
15 not). The pH also affects the ratio of polymer to protein to be used. In general, if the pH is lower, a larger excess of polymer to protein will be desired (i.e., the less reactive the N-terminal  $\alpha$ -amino group, the more polymer needed to achieve optimal conditions).  
20 If the pH is higher, the polymer:protein ratio need not be as large (i.e., more reactive groups are available, so fewer polymer molecules are needed). For purposes of the present invention, the pH will generally fall within the range of 3-9, preferably 3-6.

25 Another important consideration is the molecular weight of the polymer. In general, the higher the molecular weight of the polymer, the fewer number of polymer molecules which may be attached to the protein. Similarly, branching of the polymer should be taken into  
30 account when optimizing these parameters. Generally, the higher the molecular weight (or the more branches) the higher the polymer:protein ratio. In general, for the pegylation reactions contemplated herein, the preferred average molecular weight is about 2kDa to  
35 about 100kDa (the term "about" indicating  $\pm 1$ kDa). The preferred average molecular weight is about 5kDa to

about 50kDa, particularly preferably about 12kDa to about 25kDa. The ratio of water-soluble polymer to TRIP1 protein will generally range from 1:1 to 100:1, preferably (for polypegylation) 1:1 to 20:1 and (for  
5 monopegylation) 1:1 to 5:1.

Using the conditions indicated above, reductive alkylation will provide for selective attachment of the polymer of any TRIP1 or TP2 protein having an  $\alpha$ -amino group at the amino terminus, and  
10 provide for a substantially homogenous preparation of monopolymer/TRIP1 or TP2 protein conjugate. The term "monopolymer/TRIP1 or monopolymer/TP2 protein conjugate" is used here to mean a composition comprised of a single polymer molecule attached to a TRIP1 or TP2 protein  
15 molecule. The monopolymer/TRIP1 or monopolymer/TP2 protein conjugate preferably will have a polymer molecule located at the N-terminus, but not on lysine amino side groups. The preparation will preferably be greater than 90% monopolymer/TRIP1 or monopolymer/TP2  
20 protein conjugate, and more preferably greater than 95% monopolymer TRIP1 or TP2 protein conjugate, with the remainder of observable molecules being unreacted (*i.e.*, protein lacking the polymer moiety). These examples provide for a preparation which is at least about 90%  
25 monopolymer/ protein conjugate, and about 10% unreacted protein. The monopolymer/protein conjugate has biological activity.

For the present reductive alkylation, the reducing agent should be stable in aqueous solution and  
30 preferably be able to reduce only the Schiff base formed in the initial process of reductive alkylation. Preferred reducing agents may be selected from the group consisting of sodium borohydride, sodium cyanoborohydride, dimethylamine borane, trimethylamine  
35 borane and pyridine borane. A particularly preferred reducing agent is sodium cyanoborohydride.

Other reaction parameters, such as solvent, reaction times, temperatures, etc., and means of purification of products, can be determined based on the published information relating to derivatization of proteins with water soluble polymers.

A mixture of polymer-TRIP1 or polymer-TP2 protein conjugate molecules may be prepared by acylation and/or alkylation methods, as described above, and one may select the proportion of monopolymer/ protein conjugate to include in the mixture. Thus, where desired, a mixture of various protein with various numbers of polymer molecules attached (*i.e.*, di-, tri-, tetra-, etc.) may be prepared and combined with the monopolymer/TRIP1 or monopolymer/TP2 protein conjugate material prepared using the present methods.

Generally, conditions which may be alleviated or modulated by administration of the present polymer/TRIP1 or polymer/TP2 include those described herein for TRIP1 and TP2 molecules in general. However, the polymer/TRIP1 or polymer/TP2 molecules disclosed herein may have additional activities, enhanced or reduced activities, or other characteristics, as compared to the non-derivatized molecules.

TRIP1 or TP2 nucleic acid molecules, fragments, and/or derivatives that do not themselves encode polypeptides that are active in activity assays may be useful as hybridization probes in diagnostic assays to test, either qualitatively or quantitatively, for the presence of TRIP1 or TP2 DNA or corresponding RNA in mammalian tissue or bodily fluid samples.

TRIP1 or TP2 polypeptide fragments and/or derivatives that are not themselves active in activity assays may be useful for preparing antibodies that recognize TRIP1 or TP2 polypeptides.

The TRIP1 or TP2 polypeptides and fragments thereof, whether or not chemically modified, may be

employed alone, together, or in combination with other pharmaceutical compositions.

5 The TRIP1 and TP2 polypeptides and/or fragments thereof may be used to prepare antibodies generated by standard methods. Thus, antibodies that react with the TRIP1 or TP2 polypeptides, as well as reactive fragments of such antibodies, are also contemplated as within the scope of the present invention. The antibodies may be polyclonal,  
10 monoclonal, recombinant, chimeric, single-chain and/or bispecific. Typically, the antibody or fragment thereof will either be of human origin, or will be "humanized", i.e., prepared so as to prevent or minimize an immune reaction to the antibody when administered to a patient.  
15 The antibody fragment may be any fragment that is reactive with the TRIP1 or TP2 of the present invention, such as, Fab, Fab', etc. Also provided by this invention are the hybridomas generated by presenting TRIP1 or TP2 or a fragment thereof as an antigen to a  
20 selected mammal, followed by fusing cells (e.g., spleen cells) of the mammal with certain cancer cells to create immortalized cell lines by known techniques. The methods employed to generate such cell lines and antibodies directed against all or portions of a human  
25 TRIP1 or TP2 polypeptide of the present invention are also encompassed by this invention.

The antibodies may be used therapeutically, such as to inhibit binding of TRIP1 or TP2 to telomeres or to telomerase RNA, or to other components of the  
30 telomerase complex or proteins that bind to the telomerase complex, or to inhibit TRIP1 or TP2 activity in other ways. The antibodies may further be used for *in vivo* and *in vitro* diagnostic purposes, such as in labeled form to detect the presence of the TRIP1 or TP2  
35 in a body fluid or cell sample.

Therapeutic Compositions and Administration

Therapeutic compositions of TRIP1 or TP2 are within the scope of the present invention. Such compositions may comprise a therapeutically effective amount of a TRIP1 or TP2 polypeptide or fragment thereof (either of which may be chemically modified) in admixture with a pharmaceutically acceptable carrier. The carrier material may be water for injection, preferably supplemented with other materials common in solutions for administration to mammals. Typically, a TRIP1 therapeutic compound will be administered in the form of a composition comprising purified TRIP1 polypeptide or fragment (which may be chemically modified) in conjunction with one or more physiologically acceptable carriers, excipients, or diluents. Neutral buffered saline or saline mixed with serum albumin are exemplary appropriate carriers. Preferably, the product is formulated as a lyophilizate using appropriate excipients (e.g., sucrose). Other standard carriers, diluents, and excipients may be included as desired. Other exemplary compositions comprise Tris buffer of about pH 7.0-8.5, or acetate buffer of about pH 4.0-5.5, which may further include sorbitol or a suitable substitute therefor.

The TRIP1 or TP2 compositions can be systemically administered parenterally. Alternatively, the compositions may be administered intravenously or subcutaneously. When systemically administered, the therapeutic compositions for use in this invention may be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such pharmaceutically acceptable protein solutions, with due regard to pH, isotonicity, stability and the like, is within the skill of the art.

Therapeutic formulations of TRIP1 or TP2 compositions useful for practicing the present invention



may be prepared for storage by mixing the selected composition having the desired degree of purity with optional physiologically acceptable carriers, excipients, or stabilizers (*Remington's Pharmaceutical Sciences*, 18th edition, A.R. Gennaro, ed., Mack Publishing Company [1990]) in the form of a lyophilized cake or an aqueous solution. Acceptable carriers, excipients or stabilizers are nontoxic to recipients and are preferably inert at the dosages and concentrations employed, and include buffers such as phosphate, citrate, or other organic acids; antioxidants such as ascorbic acid; low molecular weight polypeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as Tween, Pluronic or polyethylene glycol (PEG).

The TRIP1 or TP2 composition to be used for *in vivo* administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes. Where the TRIP1 or TP2 composition is lyophilized, sterilization using these methods may be conducted either prior to, or following, lyophilization and reconstitution. The composition for parenteral administration ordinarily will be stored in lyophilized form or in solution.

Therapeutic compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

4691369  
The route of administration of the composition is in accord with known methods, e.g. oral, injection or infusion by intravenous, intraperitoneal, intracerebral (intraparenchymal), intracerebroventricular, intramuscular, intraocular, intraarterial, or intralesional routes, or by sustained release systems or implantation device which may optionally involve the use of a catheter. Where desired, the compositions may be administered continuously by infusion, bolus injection or by implantation device. Alternatively or additionally, TRIP1 or TP2 may be administered locally via implantation into the affected area of a membrane, sponge, or other appropriate material on to which TRIP1 or TP2 polypeptide has been absorbed.

15 TRIP1 or TP2 polypeptide may be administered in a sustained release formulation or preparation. Suitable examples of sustained-release preparations include semipermeable polymer matrices in the form of shaped articles, e.g. films, or microcapsules.

20 Sustained release matrices include polyesters, hydrogels, polylactides (U.S. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma ethyl-L-glutamate (Sidman *et al*, *Biopolymers*, 22: 547-556 [1983]), poly (2-hydroxyethyl-methacrylate) (Langer *et al.*, *J. Biomed. Mater. Res.*, 15: 167-277 [1981] and Langer, *Chem. Tech.*, 12: 98-105 [1982]), ethylene vinyl acetate (Langer *et al.*, *supra*) or poly-D(-)-3-hydroxybutyric acid (EP 133,988). Sustained-release compositions also may include liposomes, which can be

30 prepared by any of several methods known in the art (e.g., DE 3,218,121; Epstein *et al.*, *Proc. Natl. Acad. Sci. USA*, 82: 3688-3692 [1985]; Hwang *et al.*, *Proc. Natl. Acad. Sci. USA*, 77: 4030-4034 [1980]; EP 52,322; EP 36,676; EP 88,046; EP 143,949).

35 In other cases, TRIP1 or TP2 may be delivered through implanting into patients certain cells that have

been genetically engineered to express and secrete TRIP1 polypeptide. Such cells may be animal or human cells, and may be derived from the patient's own tissue or from another source, either human or non-human. Optionally,  
5 the cells may be immortalized. The cells may be implanted into suitable body tissues or organs of the patient.

An effective amount of the TRIP1 or TP2 composition(s) to be employed therapeutically will  
10 depend, for example, upon the therapeutic objectives such as the indication for which TRIP1 or TP2 is being used, the route of administration, and the condition of the patient. Accordingly, it will be necessary for the therapist to titer the dosage and modify the route of  
15 administration as required to obtain the optimal therapeutic effect. A typical daily dosage may range from about 0.1 µg/kg to up to 100 mg/kg or more, depending on the factors mentioned above. Typically, a clinician will administer the TRIP1 or TP2 composition  
20 until a dosage is reached that achieves the desired effect. The TRIP1 or TP2 composition may therefore be administered as a single dose, or as two or more doses (which may or may not contain the same amount of TRIP1 or TP2) over time, or as a continuous infusion via  
25 implantation device or catheter.

As further studies are conducted, information will emerge regarding appropriate dosage levels for treatment of various conditions in various patients, and the ordinary skilled worker, considering the therapeutic  
30 context, the type of disorder under treatment, the age and general health of the recipient, will be able to ascertain proper dosing.

In certain situations, it may be desirable to use gene therapy methods for administration of TRIP1 or TP2  
35 to patients suffering from HIV infection, AIDS, or other diseases for which TRIP1 or TP2 is a viable therapeutic

agent, such as, for example, premature aging and other aging disorders. In these situations, genomic DNA, cDNA, and/or synthetic DNA encoding TRIP1 or TP2, or a fragment or variant thereof, may be operably linked to a constitutive or inducible promoter (where the promoter may be homologous or heterologous) that is active in the tissue into which the composition will be injected. This construct can then be inserted into a suitable vector such as an adenovirus vector or a retrovirus vector to create a "gene therapy vector". The cells of the patient to be treated (such as, for example, T-cells in AIDS patients) can be removed from the patient, infected with the gene therapy vector using standard transfection procedures for eukaryotic cells, and tested for TRIP1 or TP2 protein production. Those cells expressing TRIP1 or TP2 can then be re-introduced into the patient.

A second method in which gene therapy can be used to modulate TRIP1 or TP2 expression is to modify the nucleotide composition of the promoter. Such modification is typically accomplished via homologous recombination methods. A DNA construct containing a portion of the TRIP1 or TP2 promoter sequence can be engineered to remove pieces of the promoter that regulate transcription. For example, the TATA box and/or the binding site of a transcriptional activator protein of the TRIP1 and/or TP2 promoter[s] may be deleted; such deletion can inhibit promoter activity thereby repressing transcription of the corresponding TRIP1 and/or TP2 gene[s]. Deletion of the TATA box or activator binding sequences in the promoter may be accomplished by generating a knockout construct comprising a mutant promoter sequence in which one or more of the TATA box and/or activator binding site nucleotides are mutated via substitution, deletion and/or insertion of one or more nucleotides. This

construct may be introduced into the appropriate cells (either *ex vivo* or *in vivo*) either directly or in a vector using standard eukaryotic transfection techniques.

5           In those situations where it is desirable to activate TRIP1 and/or TP2 expression, gene therapy and homologous recombination may be used to insert enhancer elements into the TRIP1 and/or TP2 promoters. The enhancer element(s) used will be selected based on the  
10 tissue in which one desires to activate TRIP1 or TP2; enhancer elements known to confer promoter activation in a given tissue will be selected. For example, if TRIP1 and/or TP2 are to be "turned on" in T-cells, the *lck* promoter enhancer element may be used. Here, a  
15 homologous recombination construct containing a portion of the promoter (TRIP1 and/or TP2) to be activated may be isolated, and the *lck* enhancer element may be inserted into the promoter using standard cloning techniques. The homologous recombination construct can  
20 then be introduced into the desired cells either *ex vivo* or *in vivo*.

Gene therapy methods may also be employed where it is desirable to inhibit TRIP1 or TP2 activity. Here, antisense DNA or RNA with a sequence that is  
25 complementary to: (1) full length telomerase RNA, (2) at least the portion of the telomerase RNA that interacts with TRIP1 or TP2, (3) a portion of the TRIP1 or TP2 mRNA, or (4) full length TRIP1 or TP2 mRNA can be prepared, placed into a suitable vector, and transfected  
30 into selected cells (previously removed from the patient in an *ex vivo* manner). Alternatively, the vector containing the antisense construct may be used for *in vivo* administration via microinjection, lipofection, or the like, or the desired cell types in the patient. The  
35 vector is typically selected based on its ability to

generate high levels of the anti-sense RNA in conjunction with the host cell's machinery.

Alternatively, gene therapy may be employed to create a dominant-negative inhibitor of TRIP1 or TP2.

5 In this situation, the DNA encoding a mutant full length or truncated polypeptide of TRIP1 or TP2 is inserted into a retrovirus or adenovirus, or a comparable vector, and the vector in turn is transfected into the patient's cells in either an *ex vivo* or *in vivo* manner. This  
10 TRIP1 or TP2 mutant is designed to (1) compete with endogenous TRIP1 or TP2 in forming the telomerase complex; and (2) contains one or more insertions, deletions, and/or mutations as compared to wild type TRIP1 or TP2 such that the telomerase complex is  
15 rendered functionally inactive. For example, a TRIP1 or TP2 truncation mutant in which the portion of the molecule that binds RNA (*i.e.*, approximately amino acids 1-900 of human TRIP1) is not altered, but the portion of TRIP1 such as its telomere binding domain or its  
20 protein-protein interaction domain is deleted or otherwise rendered non-functional can be generated using standard cloning techniques; this mutant can then be operably linked to a suitable promoter (one that is active in the cell type into which it will be  
25 introduced) and transfected into a patient's cells. This mutant TRIP1 protein, when over-expressed in the cells into which it is introduced, can compete with endogenous TRIP1 protein for binding of telomerase RNA and/or endogenous TP2, resulting in the formation of  
30 telomerase complexes that are inactive.

The dominant-negative expression technique has been demonstrated herein to be effective in cells transfected with mutant TP2 constructs in which the reverse transcriptase activity of TP2 has been decreased or  
35 abolished, by generating TP2 DNA constructs containing one or two point mutations at amino acid position(s) 868

0894733-101697

and/or 869, which are located in one of the reverse transcriptase domains. See the Examples section in which particular point mutations have been constructed. Other such point mutations, and/or substitution or  
5 deletion mutations which inactivate reverse transcriptase activity, and/or the TRIP1 binding domain of TP2, and/or the telomerase RNA binding domain of TP2 are also contemplated herein. Such mutant TP2 constructs can be expressed in cells which have  
10 endogenous telomerase activity and can serve to inhibit such telomerase activity by competing with native (wild type) TP2.

#### Assays to Screen for Inhibitors of TRIP1 or TP2

15 As mentioned above, it would be desirable to inhibit or significantly decrease the level of TRIP1 or TP2 activity in certain cells such as cancer cells (immortalized cells). Compounds that inhibit TRIP1 or  
20 TP2 activity could be administered either in an *ex vivo* manner, or in an *in vivo* manner by local or *iv* injection, or by oral delivery, implantation device, or the like. The assays described below provide examples of methods useful for identifying compounds that could  
25 inhibit TRIP1 activity.

For ease of reading, the following definition is used herein for describing the assays:

"Test molecule(s)" refers to the molecule(s) that is under evaluation as an inhibitor of TRIP1 or  
30 TP2, either by virtue of its potential ability to block (1) the interaction of TRIP1 or TP2 with telomerase RNA; (2) the interaction of TRIP1 or TP2 with telomere binding proteins, with the telomere itself, or with other polypeptides that comprise the telomerase complex,  
35 or (3) the active site of TRIP1 or TP2.

A. In Vitro Assays Using Purified Protein

Several types of *in vitro* assays using purified protein may be conducted to identify those compounds that disrupt telomerase activity. Such disruption may be accomplished by a compound that either inhibits the interaction of TRIP1 or TP2 with the telomeres, by a compound that inhibits TRIP1 or TP2 association with telomerase RNA or other protein components of the telomerase enzyme complex, or by a compound that blocks a reverse transcriptase motif or motifs of TP2.

In one assay, purified TRIP1 or TP2 protein or a fragment thereof (prepared for example using methods described above) can be immobilized by attachment to the bottom of the wells of a microtiter plate. Radiolabeled telomerase RNA, as well as the test molecule(s) can then be added either one at a time or simultaneously to the wells. After incubation, the wells can be washed and counted using a scintillation counter for radioactivity to determine the degree of TRIP1/telomerase RNA binding or TP2/telomerase binding RNA in the presence of the test molecule. Typically, the molecule will be tested over a range of concentrations, and a series of control "wells" lacking one or more elements of the test assays can be used for accuracy in evaluating the results. A variation of this assay involves attaching the telomerase RNA to the wells, and adding radiolabeled TRIP1 or TP2 along with the test molecule to the wells. After incubation and washing, the wells can be counted for radioactivity.

Several means other than radiolabelling are available to "mark" the TRIP1, TP2, or telomerase RNA. For example, a fusion protein of TRIP1 or TP2 wherein the DNA encoding TRIP1 is fused to the coding sequence of a peptide such as the *c-myc* epitope. TRIP1-myc



fusion protein or TP2-myc fusion protein can readily be detected with commercially available antibodies directed against *myc*.

5       Telomerase RNA can be labeled by synthesizing it with radiolabelled nucleotides such as 32-P ATP, and the level of radioactivity can then be measured by scintillation counting. Alternatively, the RNA can be labeled using biotin, digoxigenin, or a comparable compound.

10       An alternative to microtiter plate type of binding assays comprises immobilizing either TRIP1, TP2, or telomerase RNA on agarose beads, acrylic beads or other types of such inert substrates. The inert substrate containing the RNA or TRIP1 or TP2 can be  
15       placed in a solution containing the test molecule along with the complementary component (either RNA or TRIP1 or TP2) which has been radiolabeled or fluorescently labeled; after incubation, the inert substrate can be precipitated by centrifugation, and the amount of  
20       binding between TRIP1 and RNA or between TP2 and RNA can be assessed using the methods described above. Alternatively, the insert substrate complex can be immobilized in a column and the test molecule and complementary component passed over the column.  
25       Formation of the TRIP1/RNA complex or TP2/RNA complex can then be assessed using any of the techniques set forth above, *i.e.*, radiolabeling, antibody binding, or the like.

30       Another type of *in vitro* assay that is useful for identifying a molecule to inhibit TRIP1 activity is the Biacore assay system (Pharmacia, Piscataway, NJ) using a surface plasmon resonance detector system and following the manufacturer's protocol. This assay essentially involves covalent binding of either TRIP1 or  
35       telomerase RNA to a dextran-coated sensor chip which is located in a detector. The test molecule and the

complementary component can then be injected into the chamber containing the sensor chip either simultaneously or sequentially, and the amount of binding of TRIP1/RNA or TP2/RNA can be assessed based on the change in molecular mass which is physically associated with the dextran-coated side of the of the sensor chip; the change in molecular mass can be measured by the detector system.

One other assay useful for evaluating test molecule disruption of the TRIP1/RNA or TP2/RNA complex is the gel shift assay. Here, TRIP1 or TP2, telomerase RNA, and the test molecule can be incubated together. Typically, the RNA is radiolabelled using standard radioisotopes for nucleic acids (such as  $^{32}$ -P ATP). After incubation, the samples can be run on a non-denaturing acrylamide gel where the acrylamide concentration is about 4-6 percent. The migration pattern of telomerase RNA on the gel can then be evaluated. Where the TRIP1/RNA complex or TP2/RNA complex is intact during electrophoresis (even after treatment with the test molecule) migration will be slowed due to the increased molecular weight of the complex. If, however, the test molecule has sufficiently disrupted the TRIP1/RNA complex or the TP2/RNA complex, telomerase RNA will migrate in a manner comparable to control (un-treated) telomerase RNA. Migration can be detected by autoradiography.

In some cases, it may be desirable to evaluate two or more test molecules together for use in decreasing or inhibiting TRIP1 or TP2 activity. In these cases, the assays set forth above can be readily modified by adding such additional test molecule(s) either simultaneously with, or subsequently to, the first test molecule. The remainder of steps in the assay can be as set forth above.

B. In Vitro Assays Using Cultured Cells

Cultures of immortalized cells (either normal mammalian cells that have spontaneously gained the ability to replicate indefinitely, normal mammalian cells transformed with oncogenes, or mammalian cells derived from tumors) can be used to evaluate test molecules for TRIP1 or TP2 inhibition. The immortalized cells can be obtained from any mammal, but preferably will be from human or other primate, canine, or rodent sources.

In one type of cell culture assay, the immortalized cells can be cultured in standard medium such as DMEM, alpha-MEM, or RPMI. Typically, the medium would contain up to about ten percent (v:v) of fetal calf serum. Incubation is typically conducted for 1-5 days. After this incubation, the test molecule or molecules can be added, and the cells incubated for a period of 1-7 days, allowing for 3-8 cell cycles. After washing the cells to remove any residual test molecule, the cells can be harvested and telomerase activity analyzed in an *in vitro* assay such as the TRAP assay (Kim *et al*, *supra*) or the TRF assay (Harley *et al.*, 1990, *supra*). Inhibition may be manifested by a decrease in telomere length, telomerase activity, or both. For example, two known reverse transcriptase inhibitors, dideoxy GTP and AZT, have been shown to cause a decrease in telomere length in immortalized cells and a decrease in telomerase activity *in vitro* (Strahl *et al.*, *Mol. Cell. Biol.*, 16:53-65 [1996]).

In another cell assay, human immortalized cells can be transfected with a DNA construct encoding either full length TRIP1 or TP2, or a truncated version of TRIP1 or TP2. After transfection, the cells can be incubated for a period of time, after which telomerase activity can be assessed using the TRAP assay, and

telomere length assayed by the TRF or other suitable assay.

The following examples are intended for illustration purposes only, and should not be construed as limiting the scope of the invention in any way.

#### EXAMPLES

10

##### 1. Molecular Cloning of Murine TRIP1 cDNA

Standard methods for library preparation, DNA cloning, and protein expression are set forth in Sambrook *et al.*, (*Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laborite Press, Cold Spring Harbor, NY [1989]).

A cDNA library was constructed using RNA purified from adult murine colonic crypt cells. mRNA was isolated from a membrane bound polysomal fraction of RNA (Mechler *et al.*, *Meth, Enz.*, 152:241-248 [1987]). The poly(A+) mRNA fraction was isolated from the total RNA preparation using the FastTrac mRNA Isolation Kit (Invitrogen, San Diego, CA) according to the manufacturer's recommended procedure. First strand cDNA was generated by reverse transcribing the RNA using random hexanucleotides (RediPrime kit, Amersham, Arlington Heights, IL).

A random primed cDNA library was prepared from the first strand cDNA using the Superscript Plasmid System (Gibco BRL, Gaithersburg, MD). A random cDNA primer containing an internal *NotI* restriction site was used to initiate first strand synthesis and had the following double strand sequence:

35            CCTCTGCGGCCGCTACANNNNNNNNT            (SEQ ID NO: 5)  
             GGAGACGCCGGCGA'                        (SEQ ID NO: 6)

The first strand cDNA synthesis reaction was assembled using 1 µg of the mRNA and 150 ng of the *NotI* random primer. After second strand synthesis, the reaction products were extracted with the phenol:chloroform:isoamyl alcohol mixture and ethanol precipitated. The double strand (ds) cDNA products were ligated to the following ds oligonucleotide adapter (Gibco BRL):

TCGACCCACGCGTCCG (SEQ ID NO: 7)  
GGGTGCGCAGGC (SEQ ID NO: 8)

After ligation the cDNA was digested to completion with *NotI*, extracted with phenol:chloroform:isoamyl alcohol (25:24:1 ratio) and ethanol precipitated. The resuspended cDNA was then size fractionated by gel filtration using the premade columns provided with the Superscript Plasmid System (Gibco BRL) as recommended by the manufacturer. The fractions containing the largest cDNA products were ethanol precipitated and then directionally ligated into *NotI* and *SalI* digested pMOB vector DNA (Strathmann *et. al. Science* 252:802-808 [1991]). The ligated cDNA was introduced into electrocompetent XL1-Blue *E. coli* (Stratagene, LaJolla, CA) by electroporation. The library was termed cm1.

Approximately 20,000 colonies from the library were picked and arrayed into 96 well microtiter plates containing about 200 µl of L-broth, 7.5% glycerol, 50 µg/ml ampicillin and 12.5 µg/ml tetracycline. The cultures were grown overnight at 37°C, a duplicate set of microtiter plates were made using a sterile 96 pin replicating tool, and both sets were stored at -80°C for further analysis.

To sequence random cDNA clones from this library, sequencing template was prepared by PCR amplification of cloned cDNA inserts using vector primers. Glycerol stocks of cDNA clones were thawed, and small aliquots were diluted 1:25 in distilled water. Approximately 3.0 µl of diluted bacterial cultures were added to PCR reaction mixture (Boehringer-Mannheim) containing the following oligonucleotides:

10           TGTAACGACGGCCAGT           (SEQ ID NO: 9)  
              CAGGAAACAGCTATGACC       (SEQ ID NO: 10)

The reactions were incubated in a thermocycler (Perkin-Elmer 9600) with the following cycle conditions: 94°C for 2 minutes; 94°C for 5 seconds, 50°C for 5 seconds and 72°C for 3 minutes for 30 cycles and then a final extension at 72°C for 4 minutes. After incubation in the thermocycler, the reactions were diluted with about 2.0 ml of water. The amplified DNA fragments were further purified using Centricon columns (Princeton Separations) using the manufacturer's recommended procedures. In some instances, low primer and deoxynucleoside triphosphate concentrations were used in the amplification reactions, and in those instances, Centricon purification was not necessary. The PCR reaction products were sequenced on an Applied Biosystems 373A automated DNA sequencer using T3 primer:

              CAATTAACCCTCACTAAAG       (SEQ ID NO: 11)  
30

Taq dye-terminator reactions (Applied Biosystems) were conducted following the manufacturer's recommended procedures.

A search of six way translated DNA sequences from these clones was performed to isolate clones that conformed to the following criteria:

1. Potential signal peptide: Translated  
5 sequences contain the following: a methionine followed by one to three positively charged residues followed by 6-15 hydrophobic residues followed by 1-2 charged residues, followed by an open reading frame of at least residues.
- 10 2. Predicted alpha helical structure. The open reading frame contains sequences that are predicted to contain at least 30% alpha helix as assayed by the Robson/Garnier algorithm contained in the software program *Macvector 4.5*.
- 15 3. Leucine content. The open reading frame contains at least 10% Leucine residues.
4. Cysteine content. The open reading frame contains at least one but not more than 7 cysteine residues.
- 20 5. Lack of transmembrane domain. The open reading frame does not contain a sequence of 15-25 consecutive hydrophobic or uncharged residues.

One clone meeting all of these criteria, cm1-85-g3, was selected for further characterization. To  
25 identify additional sequence of this clone, a search of clones obtained from a mouse colon tissue cDNA library (prepared essentially as described above) using cm1-85-g3 as a probe resulted in the identification of clone cm3-1-e4, which had overlapping (homologous) sequence  
30 with cm1-85-g3, and contained additional 3' sequence, including a 3' termination codon. Clone cm1-85-g3 was about 1322 base pairs (bp) in length, and clone cm3-1-e4 was about 6.9 kb. To obtain the 5' portion of the coding region, PCR amplification was performed using an  
35 antisense oligonucleotide corresponding to the 5' end of the cm1-85-g3 clone and an oligonucleotide corresponding

to a portion of the pMOB vector polylinker sequence. The template for this PCR reaction was 96 DNA samples. Each sample was prepared by first plating the entire cm1 library at a density of about 10,000 clones on 96 15 cm plates. After culturing, each plate was scraped and the resultant pooled bacteria containing the clones was prepared as a glycerol stock. DNA was prepared from a portion of each pool, and 1-3  $\mu$ l of each DNA sample was then added to the individual wells.

PCR conditions were: 30 cycles, 94°C for 20 seconds; 50°C for 10 seconds, and 72°C for 30 seconds. Samples were analyzed by agarose gel electrophoresis.

A PCR fragment of about 1.5 kb was isolated from one of the PCR reactions, and was sequenced. A search of various databases with this PCR fragment resulted in the identification of a homologous sequence termed bmst2-15-g6. This clone was sequenced in its entirety, and was found to contain a methionine preceded by several stop codons, indicating a translation start site for the gene.

The three clones cm1-85-g3, cm3-1-e4 and bmst2-15-g6 overlapped to form a contiguous sequence of about 8159 bp in length. Within this sequence was an open reading frame of about 7887 bp comprising about 2629 amino acids.

A FASTA search of this open reading frame against all translated DNA sequences in the Genbank DNA Repository revealed a homology to the *Tetrahymena* telomerase P80 subunit. Several significant stretches of amino acid homology were found across this *Tetrahymena* amino acid sequence. One of these regions showed about 46 percent identity over a 90 amino acid length of the *Tetrahymena* telomerase P80 subunit. Due to its homology with *Tetrahymena* telomerase, this gene was called murine telomerase RNA interacting protein 1 ("TRIP1").



## 2. Cloning of human TRIP1 Gene

The human homolog for the murine TRIP1 gene was identified by screening a cDNA library constructed using RNA from the human colon tumor cell line LIM1863 (Whitehead *et al.*, *Cancer Res.*, 47:2704-2713 [1987]). Total RNA was isolated and the poly(A+) mRNA fraction was obtained using the FastTrac mRNA Isolation Kit (Invitrogen, San Diego, CA) according to the manufacturer's recommended procedure.

A random cDNA primer containing an internal *NotI* restriction site was used to initiate first strand synthesis. This primer had the double strand sequence as set forth above for SEQ ID NO:5 and SEQ ID NO:6. The first strand cDNA synthesis reaction was assembled using about 1µg of the mRNA and 150 ng of the *NotI* random primer (*i.e.*, SEQ ID NOS:5 and 6). A random primed cDNA library was then prepared from this first strand cDNA material using the Superscript Plasmid System (Gibco BRL, Gaithersburg, MD).

After second strand synthesis, the reaction products were extracted with the phenol:chloroform:isoamyl alcohol mixture and ethanol precipitated. The double strand (ds) cDNA products were ligated to a double strand oligonucleotide adapter with the sequence set forth above for SEQ ID NO:7 and SEQ ID NO:8.

After second strand synthesis, the reaction products were extracted with the phenol:chloroform:isoamyl alcohol mixture and ethanol precipitated. The double strand (ds) cDNA products were ligated to a double strand oligonucleotide adapter with the sequence set forth above for SEQ ID NO:7 and SEQ ID NO:8.

After ligation, the cDNA was digested to completion with *NotI*, extracted with

phenol:chloroform:isoamyl alcohol (25:24:1 ratio) and ethanol precipitated. The resuspended cDNA was then size fractionated by gel filtration using the premade columns provided with the Superscript Plasmid System (Gibco/BRL) as recommended by the manufacturer. The fractions containing the largest cDNA products were ethanol precipitated and then directionally ligated into *NotI* and *SalI* digested pSPORT vector (Gibco/BRL, Grand Island, NY). The ligated cDNA was introduced into electrocompetent XL1-Blue *E. coli* (Stratagene, LaJolla, CA) by electroporation.

The cDNA library was arrayed by plating the entire library at a density of about 10,000 clones per plate on 96 15 cm Petri plates. After incubation, each plate was scraped, and the resultant pooled bacteria was prepared as a glycerol stock. DNA was prepared from an aliquot of each pool, digested with *NotI*, electrophoresed on a 1% agarose gel and transferred to a charged nylon membrane for Southern blotting. Each of the 96 lanes on the gel thus contained about 10,000 cDNA clones. An approximately 500bp *BamHI/HindIII* fragment of clone cm1-85-g3 was random prime labeled using standard methods and hybridized to the Southern blot. Hybridization was conducted at 50°C for at least two hours using Rapid Hyb buffer (Amersham, Arlington Heights, IL) and following the manufacturer's protocol. About ten percent of the samples hybridized to the probe. Lanes corresponding to DNA pools 54, 58 and 87 contained the largest inserts, and so these were selected for further analysis.

Glycerol stocks of bacteria containing the indicated pooled clones were plated directly on to nitrocellulose filters covering agar plates, grown for several hours at 30°C, lysed, and hybridized to the cm1-85-g3 500 bp random primed probe. Hybridization conditions were as above using Rapid Hyb buffer.

Positive clones were picked and rescreened to isolate single clones from each stock. The three selected clones, called 54, 58, and 87, contained significant overlapping sequence with each other. To identify additional 5' sequence for the human TRIP1 gene, the largest of the three clones, clone 54, was used to generate one antisense oligonucleotide positioned near its 5' end for a PCR primer. The second PCR primer corresponded to the pSPORT vector. The templates for PCR were the same 96 well pools described above. PCR conditions were: 30 cycles, 94°C for 20 seconds; 50°C for 10 seconds, and 72°C for 30 seconds. Samples were analyzed by agarose gel electrophoresis using the antisense oligonucleotide together with an oligonucleotide sequence found in the pSPORT polylinker.

An approximately 1.5 kbp band was identified in pool 96. This pool was then plated and screened as above except that the filters were hybridized at 60°C using Rapid Hyb buffer as above for at least two hours. The probe was an antisense oligonucleotide to the 5' end of clone 54, and was radiolabeled at the 5' end using standard methods as follows. About 170 ng of the probe was incubated at about 37°C for about one hour in a solution containing about 200 µCi of 32-P labeled ATP (Amersham, Arlington Heights, IL) and about 20 U of Polynucleotide Kinase (Boehringer Mannheim, Indianapolis, IN), using a buffer provided by the manufacturer. Radiolabeled oligonucleotide was separated from unincorporated nucleotide by centrifugation through a G25 Quickspin column (Boehringer Mannheim) according to the manufacturer's protocol.

To identify the 3' region of the human TRIP1 gene, a sense oligonucleotide corresponding to the 3' end of clone 54 and an oligonucleotide sequence corresponding to the pSPORT polylinker were used in a

PCR reaction. The same 96 well pools were used as a template for PCR reactions.

PCR conditions were: 30 cycles, 94°C for 20 seconds; 55°C for 10 seconds, and 72°C for 30 seconds.

5 Samples were analyzed by agarose gel electrophoresis.

A 3 kb PCR product was identified from DNA pool 63. This pool was then plated and screened as above. The probe for this reaction was a sense oligonucleotide to the 3' end of clone 54 which was  
10 radiolabeled at the 5' end using standard methods. Two colonies containing DNA clones which strongly hybridized to the probe were identified then sequenced in their entirety. These clones were termed 96 and 63.

To identify the remaining 3' portion of the  
15 coding sequence, another round of PCR was conducted. Here, the primers used were (1) a sense oligonucleotide to the 3' end of clone 63 and (2) an oligonucleotide corresponding to the SP6 of the pSPORT vector. PCR conditions were: 30 cycles, 94°C for 20 seconds; 55°C  
20 for 10 seconds, and 72°C for 30 seconds. The templates for PCR were the same 96 well pools. Samples were analyzed by agarose gel electrophoresis. An approximately 200 bp fragment was identified in pool 15. This pool was then plated and screened as above by  
25 hybridizing the filters with a radiolabeled probe. The probe for this reaction was a sense oligonucleotide to the 3' end of clone 63 which was radiolabeled at the 5' end using standard methods. This clone, clone 15, was sequenced in its entirety and was found to possess a  
30 termination codon.

### 3. Murine TRIP1 Protein Preparation

A truncated version of murine TRIP1 protein encoding amino acids 1-871 was prepared as follows. The  
35 DNA encoding this region was obtained by PCR using the following two oligonucleotides: (1) an oligonucleotide

encoding a *SalI* restriction site followed by the first six amino acids of murine TRIP1 and (2) an oligonucleotide corresponding to amino acids 866-871 followed by a TAG stop codon and a *SalI* restriction site. The template for this reaction were clones cm1-85-g3, cm3-1-e4 and bmst2-15-g6. PCR reactions were 15 cycles, 94°C for 20 seconds, 55°C for 10 seconds, and 72°C for 30 seconds.

This reaction resulted in a band of approximately 2.6 kb on an agarose gel. This band was purified from the gel, digested with *SalI* and cloned into the *XhoI* site of the vector pCR3MycTag. pCR3MycTag was prepared as follows. The vector pCR3 (Invitrogen, San Diego, CA) was digested with *KpnI* and *XhoI*. A nucleic acid molecule encoding two copies of the *c-myc* epitope and an initiation Methionine was inserted into pCR3. The sequence of this insert is set forth below as SEQ ID NO:12. The resulting plasmid containing the TRIP1 insert (cDNA encoding amino acids 1-871) was termed pCR3MycTag2.

GGTACCGCCAGCCGAGCCACATCGCTCAGACACCATGATCGCAAATGTGAATATTG  
CTCAGGAACAAAAGCTTATTTCTGAAGAAGACTTGGCTCAGGAACAAAAGCTTATT  
TCTGAAGAAGACTTGGCTCAGCAGAGTGGCGGAGGACTCGAG (SEQ ID  
NO:12)

A second plasmid, pCR3MycTag3, which contained the cDNA encoding full length murine TRIP1, was prepared as follows. The plasmid pCR3MycTag2 was digested with *EcoRI* and *XbaI* (which served to delete the cDNA encoding amino acids 816-871 from the vector), and an *XbaI/SalI* linker was ligated into the digested plasmid. An *EcoRI/SalI* fragment 5.4 of clone cm3-1-e4 (corresponding to amino acids 816 to 2627 of murine TRIP1) was ligated into the vector. The resulting plasmid, pCR3MycTag3, has the following components (from 5' to 3'): an

initiation codon, two *c-myc* epitopes, and the full length murine TRIP1 cDNA.

Full length and truncated (amino acids 1-871) murine TRIP1 protein was prepared as follows. Plasmid DNA from pCR3MycTag2 and pCR3MycTag3 was transfected into murine neuroblastoma N2A cells (American Type Culture Collection, catalog no. CCL131) by lipofection using the Perfect Lipid Transfection kit (Invitrogen, San Diego, CA). These cells are commonly used for transient and stable expression of foreign proteins. About 24 hours prior to transfection, the cells were seeded at about 700,000 per 100 mm dish in DMEM plus ten percent fetal calf serum, and PSG (penicillin, streptomycin, and glutamine). For lipofection, the cells were placed in about 6 ml of Optimem I reduced serum medium (Gibco/BRL, Grand Island, NY) and about 174 µg of Pfx-6 (Invitrogen) and 29 µg of DNA were added. The cells were incubated for about 4 hours after which time the medium was replaced with fresh DMEM, fetal calf serum, and PSG medium as described above. The cells were harvested after about 24 hours, and were lysed using a Qiagen shredder (Qiagen, Chatsworth, CA) according to the manufacturer's protocol. Protein lysates were electrophoresed by 6 percent SDS-PAGE, transferred to a nylon membrane using standard methods, and incubated with a mouse monoclonal anti-myc antibody (Oncogene Research Products, Cambridge, MA). Binding of the anti-myc antibody was detected with a HRP-conjugated secondary antibody, and the complex was visualized using ECL (Amersham, Arlington Heights, IL) following the manufacturer's protocol. Cells transfected with the vector containing the TRIP1 truncated cDNA showed a prominent band of about 97 kD (corresponding to a polypeptide of about 871 amino acids), while cells transfected with the vector containing full length TRIP1 showed a prominent band of about 280 kD (corresponding

to a polypeptide of about 2625 amino acids). These results indicated that TRIP1 truncated or full length protein was expressed in the cells.

5 4. Murine TRIP1 RNA-Binding Assay

To determine whether mTRIP1 had a specific interaction with the RNA molecule known to be mouse telomerase RNA, the three hybrid assay as described by SenGupta *et al.* (*Proc. Natl. Acad. Sci USA*, 93:8496-8501 [1996]) was used. The starting plasmid described by SenGupta *et al.*, pMS2-2, was altered by inserting, using standard ligation methods, a DNA encoding the full length mouse telomerase RNA transcript (mTR; Blasco *et al.*, *Science*, 269:1267-1270 [1995]) into the *SmaI* polylinker site of pMS2-2 in the same orientation as the two MS2 DNA sequences at the 3' end of the polylinker region. (The RNA molecules  $\alpha$ -mTR, *TLC1*, IRE and the mutant mTR molecules, all described in Table I below, were constructed in this same manner; U2, U4, and U6 were similarly tagged with the MS2 hairpins, but were inserted into a different URA3 selectable yeast plasmid, pRS316 [Sikorski *et al.*, *Genetics*, 122:19-27, 1989]).

After this ligation, the resultant plasmid was digested with *EcoRI*, and the approximately 700 base pair (bp) fragment containing 5' to 3', mTR and the two MS2 DNA sequences, was isolated by standard agarose gel purification methods. This 700 bp fragment was then inserted into plasmid pIIIEx426 (SenGupta *et al.*, *supra*) which had been previously digested with *EcoRI*. This plasmid was referred to as pIII-mTR.

A second plasmid was also prepared as follows. The starting plasmid was pACTII (Legrain *et al.*, *Nuc. Acids Res.*, 22:3241-3242 [1994]). pACTII was first digested with the enzyme *BamHI*, and the ends were blunted using T4 DNA polymerase. An *SspI/XbaI* fragment of plasmid pCR3MycTag2 (see above) was isolated using

standard gel purification methods and blunt ended using T4 DNA polymerase. This fragment, which was about 2739 bp, contained 126 bp (42 amino acids) of vector sequence at the 5' end and the first 871 amino acids of mTRIP1.

- 5 The fragment was inserted into the *Bam*HI digested pACTII, and the resultant plasmid was referred to as pACTII/MTRIP1-S/X.

- 10 Plasmids pACTII/MTRIP1-S/X and pIII-mTR were introduced into yeast cells (strain L40-coat; SenGupta *et al.*, *supra*) which had been cultured in standard yeast media (YEPD; Sherman *et al.*, *Meth. Yeast Genet.*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY [1983]). Introduction (also referred to as transformation) of the plasmids was accomplished using
- 15 standard methods such as those described by Chen *et al.* (*Curr. Genet.*, 21:83-84 [1992]). Co-transformants (*i.e.*, those yeast cells that contained both introduced plasmids) were selected by culturing the cells on yeast agar plates lacking leucine and uracil (SD-ura-leu;
- 20 Sherman *et al.*, *supra*) for two days at about 30°C. Eight separate, randomly selected colonies of cells that grew on these plates were repatched on fresh SD-ura-leu plates, and incubated as above. A small portion of each colony was plated on to yeast agar plates lacking
- 25 uracil, leucine, histidine, and containing 5-20 mM 3-aminotriazole (Sigma, ST. Louis, MO), and the plates were incubated 3 days at about 30°C, after which time the number of colonies that grew (out of a total of eight) was assessed.



The results are shown in Table II below.

Table II

RNA	PROTEIN	INTERACTION		
		5mM	10mM	20mM
mTR	mTRIP1	8/8	8/8	8/8
mTR-1	mTRIP1	8/8	8/8	8/8
mTR-3	mTRIP1	8/8	8/8	8/8
mTR-27	mTRIP1	8/8	8/8	8/8
U2	mTRIP1	0/8	0/8	0/8
U4	mTRIP1	0/8	0/8	0/8
U6	mTRIP1	0/8	0/8	0/8
<i>TLC1</i>	mTRIP1	0/8	0/8	0/8
$\alpha$ -mTR	mTRIP1	0/8	0/8	0/8
mTR-1	IRP	0/8	0/8	0/8
IRE	IRP	8/8	8/8	8/8
IRE	mTRIP1	7/8	7/8	7/8
MS2	mTRIP1	7/8	6/8	5/8

5

In Table II, the column labeled "RNA" refers to the MS2 tagged RNA molecules that were tested. mTR is wild type mouse telomerase RNA; mTR-1 is a substitution mutant of mTR and contains a T instead of a C at position 142 (relative to the transcription start site; see Blasco *et al. supra*), a C instead of a G at position 202, and an A instead of a G at position 227; mTR-3 contains an A instead of a G at position 272 and is also an insertion mutant of mTR in which two nucleotides, A and G, were inserted after nucleotide 268 in the mTR transcript (Blasco *et al., supra*); mTR-27 is a substitution mutant of mTR that contains an A instead of a G at position 33; U2, U4, and U6 are snRNAs (Ares, *Cell*, 47:44-59 [1986]; Tollervey *et al., Cell*, 35:753-762 [1983]; Brow *et al., Nature*, 334:213-218 [1988]);

20

*TLC1* is the yeast telomerase RNA gene (Singer *et al.*, *Science*, 266:404-409 [1994]);  $\alpha$ -mTR is the mTR sequence cloned in the anti-sense direction relative to the MS2 hairpins; IRE is the rat iron regulatory element RNA (Fields *et al.*, *supra*); and MS2 refers to the MS2 hairpins without additional RNA attached.

The column labeled "Protein" refers to proteins that were co-introduced along with the test RNA molecules to evaluate RNA-protein interaction in the three hybrid assay. "mTRIP1" is the amino terminal fragment of the mTRIP1 gene and consists of the amino terminal 871 amino acids of the protein; and IRP is the iron regulatory element binding protein (SenGupta *et al.*, *supra*).

The column labeled "Interaction" refers to the concentration (5, 10, or 20 mM) of 3-aminotriazole on the yeast agar plates.

The number of colonies out of a total of eight that showed detectable growth after 3 days is indicated for each RNA/protein pair. As can be seen, the mouse telomerase RNA, whether wild type or mutant, specifically interacted with mTRIP1. With the exception of IRE, the other RNA molecules, U2, U4, U6, *TLC1*, and  $\alpha$ -mTR, did not interact with mTRIP1. MS2 alone interacted with mTRIP1 to some degree at low concentrations of 3-aminotriazole. Specificity of binding of mTR was further confirmed by demonstrating that IRP, which is known to interact with IRE (and was therefore used as a positive control), did not interact with mTR-1.

## 5. Cloning of Human TP2 Gene

The strategy for cloning the human TP2 gene was similar to that used for cloning the human TRIP1 gene.

The human TP2 cDNA was identified by screening a cDNA library constructed using RNA from the human colon tumor cell line LIM1863 (Whitehead *et al.*, *Cancer Res.*, 47:2704-2713 [1987]). Total RNA was isolated and the poly(A+) mRNA fraction was obtained using the FastTrac mRNA Isolation Kit (Invitrogen, San Diego, CA) according to the manufacturer's recommended procedure.

A random cDNA primer containing an internal *NotI* restriction site was used to initiate first strand synthesis. This primer had the double strand sequence as set forth above for SEQ ID NO:5 and SEQ ID NO:6. The first strand cDNA synthesis reaction was assembled using about 1µg of the mRNA and 150 ng of the *NotI* random primer (i.e., SEQ ID NOS:5 and 6). A random primed cDNA library was then prepared from this first strand cDNA material using the Superscript Plasmid System (Gibco BRL, Gaithersburg, MD).

After second strand synthesis, the reaction products were extracted with the phenol:chloroform:isoamyl alcohol mixture and ethanol precipitated. The double strand (ds) cDNA products were ligated to a double strand oligonucleotide adapter with the sequence set forth above for SEQ ID NO:7 and SEQ ID NO:8.

After ligation, the cDNA was digested to completion with *NotI*, extracted with phenol:chloroform:isoamyl alcohol (25:24:1 ratio) and ethanol precipitated. The resuspended cDNA was then size fractionated by gel filtration using the premade columns provided with the Superscript Plasmid System (Gibco/BRL) as recommended by the manufacturer. The fractions containing the largest cDNA products were ethanol precipitated and then directionally ligated into *NotI* and *SalI* digested pSPORT vector (Gibco/BRL, Grand Island, NY). The ligated cDNA was introduced into

electrocompetent XL1-Blue *E. coli* (Stratagene, LaJolla, CA) by electroporation.

The cDNA library was arrayed by plating the entire library at a density of about 10,000 clones per plate on 96 15 cm Petri plates. After incubation, each plate was scraped, and the resultant pooled bacteria was prepared as a glycerol stock. Plasmid DNA was prepared from each stock using standard alkaline lysis procedures, and each stock DNA was then used as a template for PCR in order to identify those stocks containing TP2 cDNA sequences. The PCR reactions contained approximately 25 pmol of primer, 50-220 ng of template, both in Boehringer PCR Reaction buffer, to which was then added about 1.25U TAQ polymerase (Boehringer) in a volume of about 25  $\mu$ l. PCR conditions were 30 cycles, 94°C for 20 seconds; 50°C for 10 seconds, and 72°C for 30 seconds. The primers used for PCR were:

CCAAGTTCCTGCACTGGCTGAT (SEQ ID NO:15)

GCTCGTAGTTGAGCACGCTGAA (SEQ ID NO:16)

Samples were analyzed by agarose gel electrophoresis. Three stocks had a PCR band on the gel corresponding to the expected size of about 380 nucleotides. These stocks were selected for further analysis as follows.

Glycerol bacteria stocks of the three positive pools were plated directly on to Magnalift brand nylon membranes (Micron Separations, Westborough, MA) covering agar plates, grown for several hours at 30°C, and lysed in 0.5 M NaOH and 1.5 M NaCl for 7 minutes. The membranes were then neutralized in 1M Tris-HCl, pH 8.0.

and then baked at about 80°C for at least 2 hours in a vacuum oven. The membranes were then subjected to Proteinase K lysis by incubating them in a solution of 0.1M Tris-HCl, pH 8.0, 0.15M NaCl, 10 mM EDTA, 0.2 percent SDS, and about 50 µg/ml of Proteinase K (Boehringer).

Hybridization of the membranes was conducted for about two hours at about 60°C using Rapid Hyb buffer (Amersham, Arlington Heights, IL) following the manufacturer's protocol. The probe for this hybridization consisted of the primers used for PCR above (*i.e.*, SEQ ID NOS:15 and 16). About 170 ng total of the probe mixture was radiolabeled prior to hybridization as follows: About 170 ng of the probe was incubated at about 37°C for about one hour in a solution containing about 200 µCi of 32-P labeled ATP (Amersham, Arlington Heights, IL) and about 20 U of Polynucleotide Kinase (Boehringer Mannheim, Indianapolis, IN), using a buffer provided by the manufacturer. Radiolabeled oligonucleotide was separated from unincorporated nucleotide by centrifugation through a G25 Quickspin column (Boehringer Mannheim) according to the manufacturer's protocol. The results of this screen yielded one positive clone from the three pools screened. This clone, called #32, was sequenced using standard methods and was found to be about 2859 base pairs in size and was believed to be lacking both the 5' and 3' ends. Based on some nucleic acid homology to the telomerase polypeptides reported by Lingner *et al.* (*Science*, [1997], *supra*), it was determined that clone #32 was likely a second human telomerase subunit. The nucleic acid sequence of clone 32, referred to as a partial sequence of "telomerase protein 2" or "TP2" is shown in Figure 5 (SEQ ID NO:13), and the translated amino acid sequence is shown in Figure 6 (SEQ ID NO:14). Seven reverse transcriptase motifs (based on the

information on reverse transcriptases as set forth by Xiong *et al*, *supra*,) are present in TP2, suggesting that this protein contains reverse transcriptase activity. These motifs are present in the region of nucleotides  
5 1920-2820.

In addition to the reverse transcriptase motifs found in TP2, other significant regions of this protein include the amino acid sequences FFYVTE (SEQ ID NO: 17), RFIPK (SEQ ID NO:42), GIPQGS (SEQ ID NO:43) and  
10 LLLRLVDDFLL (SEQ ID NO:44).

To identify the 3' end of the TP2 gene, an oligonucleotide corresponding to a region near the 3' end of clone 32 was prepared. The sequence of this oligonucleotide was:

15 TGGATGATTTCTTGTTGGTGACAC (SEQ ID NO: 21)

This oligonucleotide was used for PCR in combination with an oligonucleotide that hybridizes to the SP6 viral  
20 RNA transcription start site of the pSPORT vector. Thirty three cycles of PCR were conducted on all pools of the library under the following conditions: 94°C for about 15 seconds; 62°C for about 15 seconds; and 72°C for about 30 seconds. The PCR products were evaluated  
25 by agarose gel electrophoresis and about seven positives were obtained. Further analysis of these clones indicated that they were not TP2 cDNAs. Therefore, a second screening approach was used.

Clone #32 was digested with the restriction  
30 endonuclease *MluI* to generate two fragments, a smaller 5' fragment, and a larger 3' fragment. This larger fragment, was digested with *XhoI* to generate 5' and 3' fragments. The 3' fragment, which was about 830 base pairs, was used as a probe to screen all pools of the  
35 cDNA library. The probe was labeled using the standard random primer label technique. The library was prepared

for screening as described above, and the filters were hybridized with the probe for about 2 hours at about 60°C in Rapid Hyb buffer (Amersham, Arlington Heights, IL) after which the filters were washed under stringent conditions. Seven positives were identified, and these were subjected to PCR analysis using the primer SEQ ID NO:17 and the SP6 primer. PCR conditions were as described above for these primers. Three positives were identified when PCR products were evaluated by agarose gel electrophoresis, however, upon sequencing, none of the positives contained additional TP2 sequence as compared with clone 32.

The cDNA library was rescreened using the same probe (the 830 base pair *XhoI* fragment) under the same conditions as set forth above. About eight positives were obtained, and these positives were re-plated in order to isolate single clones from each of the eight pools. After plating, the cells were grown up, and the plasmid DNA was isolated using standard miniprep procedures. Four of the plasmid DNA clones were sequenced. One clone, TP2-15 was about 1.1 kb in size, 133 bases of which overlapped with the 3' end of clone #32. The remaining 949 bases comprised new TP2 sequence at the 3' end, and also contained a stop codon. The DNA sequence of these additional 949 bases is set forth in Figure 7. The full length TP2 gene, which comprises clone 32 plus the 949 bases of TP2-15, is set forth in Figure 8, and the putative amino acid sequence of full length TP2 is set forth in Figure 9.

Single or double point mutations were made in the reverse transcriptase domain of full length TP2 at position 868 (D to A; referred to as the "5-1" mutation); 869 (D to A, referred to as the "5-2" mutation); and 868 and 869 (both, D to A; referred to as the "5-1,2" mutation). These point mutations were prepared as follows:

First, a TP2 construct containing the marker sequence FLAG (DYKDDDDK; SEQ ID NO: 22) was prepared by synthesizing a both a sense and an antisense oligonucleotide containing (for the sense strand), from 5' to 3', the restriction enzyme sequence for *HindIII*, an ATG start codon, the DNA sequence encoding the FLAG peptide sequence, and the restriction enzyme sequence for *EcoRI*. The sense oligonucleotide has the sequence:

10 AGCTTGGTACCAACATGGACTACAAGGACGACGATG (SEQ ID NO:23)

The antisense oligonucleotide has the sequence:

15 AATTCCTTGTTCATCGTCGTCCTTGTAGTCCATGTT (SEQ ID NO:24)

The two oligonucleotides were annealed by heating them together at 95°C, and then cooling the mixture slowly to room temperature. The resulting double stranded oligonucleotide was inserted into the *HindIII* and *EcoRI* sites of the pCR3 vector (Invitrogen, Inc., Carlsbad, CA).

TP2 clone 32 was excised from the pSPORT vector using restriction enzymes *EcoRI* and *NotI*, and the resulting fragment was inserted into the same sites of the pCR3 FLAG vector. Clone 15, containing the 3' end of TP2 (see Example 5 above) was digested with restriction enzymes *BamHI* and *XbaI*, and inserted into the same sites of the pCR3 FLAG/clone32 vector, resulting in a vector that expresses the full length human TP2 protein with a FLAG peptide located at the amino terminus of the TP2 protein.

Wild type and mutant TP2 proteins were next generated using the PCR mutagenesis strategy outlined in Figure 10. Six individual PCR reactions were conducted



in order to generate a series of small fragments, termed "primary PCR products" that were subsequently amplified to make the final constructs. First, six sets of two PCR primers were used to conduct six PCR reactions to amplify particular regions of TP2 and incorporate the desired point mutation(s). The regions of TP2 amplified by each of the six primer pairs are shown in Figure 10 and are labeled as reactions 1-6.

The PCR reactions were carried out using the following sets of primers. For reactions 2, 4, and 6, the 5' primer was:

CGTTTGGTGGCTGATTTCTTGTTGGTGAC (SEQ ID NO: 25)

and the 3' primers were, respectively:

GTCACCAACAAGAAATCAGCCACCAAACG (SEQ ID NO: 26)

GTCACCAACAAGAAAGCATCCACCAAACG (SEQ ID NO: 27)

GTCACCAACAAGAAAGCAGCCACCAAACG (SEQ ID NO: 28)

For reactions 1, 3, and 5, the 3' primer was:

GAATTCTAGATCACTTGTCATCGTCCTTGTTAGTCGTCAGGATGGTCTTGAAGTC (SEQ ID NO: 29)

This primer included 3' sequence of TP2, together with a sequence encoding FLAG peptide, followed by an *XbaI* restriction site.

The 5' primers for reactions 1, 3, and 5 were, respectively:

CGTTTGGTGGCTGATTTCTTGTTGGTGAC (SEQ ID NO: 30)

CGTTTGGTGGATGCTTTCTTGTTGGTGAC (SEQ ID NO: 31)

CGTTTGGTGGCTGCTTTCTTGTTGGTGAC (SEQ ID NO: 32)

5           The template used in reactions 1-6 was the  
pCR3 FLAG TP2. Each reaction was carried out using PFU  
polymerase (Stratagene, La Jolla, CA) according to the  
manufacturer's instructions. Thirty cycles were  
conducted for each reaction, each cycle consisting of  
10 96°C for 15 seconds, 62°C for 15 seconds, and then 72°C  
for 2 minutes. PCR products were extracted from an  
agarose gel.

          To prepare the final mutant constructs used  
for generating full length mutant TP2 cDNA molecules,  
15 PCR products 1 and 2 (containing the 5.1 mutation) were  
used as templates in PCR reaction 7 in which the 5'  
primer for the reaction was SEQ ID NO:25, and the 3'  
primer for the reaction was SEQ ID NO:29. Conditions  
for PCR were identical to those described immediately  
20 above, except that the 72°C step was carried out for 4  
minutes. The same strategy was used to generate a  
mutant construct in reaction 8, where PCR products 3 and  
4 (containing the 5.2 mutation) were used as template,  
SEQ ID NO: 25 was the 5' primer for the reaction, and  
25 SEQ ID NO: 29 was the 3' primer for the reaction.  
Similarly, PCR products 5 and 6 (containing the 5.1 and  
5.2 double mutation) were used as template for PCR  
reaction number 9, where SEQ ID NO: 25 was used as the  
5' primer, and SEQ ID NO: 29 was used as the 3' primer.

30           The resultant final four PCR products  
(reactions 7, 8 and 9 in Figure 10, together with the  
wild type PCR fragment) were digested with restriction  
enzymes *Bam*HI and *Xba*I, and the fragments were cloned  
into sites of the pCR3 FLAG 32 vector previously cut  
35 with enzymes *Bam*HI and *Xba*I. This resulted in the  
generation of four TP2 expression vectors containing a

FLAG tag at the 3' end. These vectors had identical inserts except for the specified mutated residues in positions at D868 and/or D869. The coding region of each vector was sequenced using an Applied Biosystems  
5 sequencer according to the manufacturer's directions.

#### 6. TP2 Antibody Preparation

Anti-TP2 rabbit polyclonal antiserum was  
10 prepared by injecting rabbits with the following TP2 peptide:

SEAEVRQHREARPALLTSRLRFIPKC (SEQ ID NO:33)

15 This peptide is also referred to herein as "TP2 specific peptide". Prior to injection, the peptide was coupled to the inert protein KLH, purified, and injected into rabbits. About one month later, the injection was repeated. Two weeks later, blood was drawn from the  
20 rabbits and analyzed for anti-TP2 antibodies by Western blot analysis of cell lysate from cells previously transfected with TP2 DNA. Upon confirmation of antibody production, antiserum produced by the rabbits was then collected using standard procedures, and affinity  
25 purified by passing it over an affinity column consisting of TP2 peptide covalently bound to Affigel™ beads (Biorad Corp., Richmond, CA). Antibody was eluted from the column using low pH buffer.

In preparation for the immunoprecipitation  
30 experiments, for each immunoprecipitation reaction, about 5 µg affinity purified antiserum was bound to about 10 µL of protein G Sepharose beads (Sigma Chemical Co., St. Louis, MO).

## 7. Biological Activity of TP2

To assess the association of TP2 with telomerase activity, three mammalian cell lines, HeLa cells, transfected mouse neuroblastoma N2A cells, and transfected human embryonic kidney 293 cells, were evaluated.

### A. Detection of TP2 in HeLa Cells

HeLa cells (Zhou *et al.*, *Genes and Devel.*, 6:1964-1974 [1992]) were cultured in standard Dulbecco's modified Eagle's medium, grown to sub-confluence, after which cell lysate was prepared as described by Prowse *et al* (*Proc. Natl. Acad. Sci., USA*, 90:1493-1497 [1993]).

Briefly, cells were pelleted, washed, and lysed in a hypotonic buffer using a Dounce homogenizer. The cell extract was centrifuged at about 100,000 X g for about one hour, and the supernatant, termed "S100", was removed and frozen in aliquots at minus 80°C.

Two assays were used to evaluate the role of TP2 in telomerase activity; these included (1) the telomerase bioassay (Kim *et al.*, *supra* [1994]), and (2) an immunoprecipitation assay.

To evaluate whether TP2 is essential for telomerase activity, HeLa cell lysates were divided into aliquots, and some aliquots were incubated with various antibodies prior to conducting the telomerase activity assay. Polyclonal TP2 or control antibodies (all of which were previously attached to protein G sepharose beads at a proportion of about 5 µg antibody to about 10µl of protein G) were incubated with about 1 mg of HeLa cell lysate for about 1 hour at about 4°C.

In some cases, as indicated below, the antibody was first incubated with a peptide (either TP2 specific or non-specific). Non-specific peptides, termed

P1 and P3, were derived from the TRIP1 amino acid sequence.

The P1 sequence is:

5

RSKRRSRQPPRPQKTERPFSEGRK

(SEQ ID NO: 34)

The P3 sequence is:

10

DPDASGTFRSCPPEALKDL

(SEQ ID NO: 35)

Specific (TP2) peptide is that which is set forth above for preparation of the TP2 antiserum. Peptide-antibody incubations were at room temperature for about 30 minutes, after which the HeLa cell lysate incubations were conducted as described immediately above. The antibodies added to various cell lysate aliquots were:

- 1) no added antibody
- 2) control antibody (anti-Myc; Pharmingen, San Diego, CA)
- 3) control antibody (anti-GST; Upstate Biotechnology, Inc., Lake Placid, NY)
- 4) anti TP2 peptide antiserum
- 5) anti TP2 peptide antiserum previously incubated with about 30  $\mu$ g non-specific peptide
- 6) anti-TP2 peptide antiserum previously incubated with about 60  $\mu$ g non-specific peptide
- 7) anti-TP2 peptide antiserum previously incubated with about 30  $\mu$ g TP2-specific peptide

8) anti-TP2 peptide antiserum previously incubated with about 60  $\mu$ g TP2-specific peptide.

5           After incubation, cell lysates were  
centrifuged to pellet the Sepharose beads (containing  
antibody and immunoprecipitate). The immunoprecipitates  
were washed with a hypotonic buffer containing 0.1 M  
NaCl (Prowse *et al.*, *supra*), and both the supernatants  
10 (approximately 5  $\mu$ g total protein) and  
immunoprecipitates (about 2  $\mu$ L protein G beads) were  
then tested for telomerase activity using the method  
described by Kim *et al.*, *supra*. Briefly, this method  
involves incubating an aliquot of telomerase extract or  
15 immunoprecipitate with a substrate oligonucleotide,  
termed a TS oligonucleotide, in the presence of  $^{32}$ P  
labeled and unlabelled deoxynucleotides and appropriate  
buffer conditions. Following elongation of the TS  
oligonucleotide by telomerase (assuming telomerase is  
20 active in the extract or immunoprecipitate), Taq  
polymerase and an oligonucleotide (termed "CX") for the  
amplification of the telomeric repeats is added, and PCR  
amplification is performed. The products can then be  
resolved by electrophoresis of the products on a non-  
25 denaturing acrylamide gel. After electrophoresis, the  
gel is dried and visualized using a phosphoimager.

Prior to the telomerase assay, some of the  
immunoprecipitates or cell extracts were incubated in  
the presence of ribonuclease A, referred to throughout  
30 as "RNase". Where RNase was used, about 1  $\mu$ g of RNase  
(Sigma Chemical Co., St. Louis, MO) was incubated with  
either about 5  $\mu$ g of HeLa cell lysate, or about 2-3  $\mu$ L  
of protein G immunoprecipitate. Incubations were  
conducted for about 5 minutes at room temperature.

35           To determine whether the immunoprecipitates  
contained TP2, Western blot analysis was conducted on an

aliquot of each immunoprecipitate as follows. The immunoprecipitate remaining in each aliquot after the activity assay was subjected to SDS-PAGE and then transferred to a PVDF membrane. For detection of TP2 on the Western blot, the blot was first incubated in TBST buffer (Tris-buffered saline in 0.5 percent Tween-20 and 5 percent dried milk) for about 1 hour at room temperature, and then incubated with 0.2  $\mu$ g/ml of anti-TP2 antibody, followed by horseradish peroxidase conjugated secondary antibody (Amersham, Arlington Heights, IL). The blot was visualized using an ECL kit (Amersham, Arlington Heights, IL).

The results of the telomerase activity assay and the Western blot analysis are shown in Figure 11A-C. Figure 11A is a 12 percent acrylamide gel showing the results of telomerase activity assay for the cell lysates (Lanes 1-2) and supernatants after immunoprecipitation (Lanes 3-9). Lane 1 represents control HeLa cell lysate without antibody added and without immunoprecipitation. Lane 2 is the same as Lane 1, except that RNase was added prior to the telomerase assay. Lanes 3 and 4 represent the supernatants of cell lysates which were preincubated with either anti-Myc antibody (Lane 3) or anti-GST antibody (Lane 4) prior to telomerase activity assays and immunoprecipitation. Lane 5 represents the supernatant of cell lysate which was pre-incubated with anti-TP2 antibody and then immunoprecipitated; Lanes 6 and 7 represent supernatants of cell lysates that were pre-incubated with anti-TP2 antibody in the presence of 30  $\mu$ g (Lane 6) or 60  $\mu$ g (Lane 7) of non-specific peptide 3. Lanes 8 and 9 represent supernatants of cell lysates that were pre-incubated with anti-TP2 antibody in the presence of 30  $\mu$ g (Lane 8) or 60  $\mu$ g (Lane 9) of TP2 peptide which recognizes the anti-TP2 antibody. As can be seen in

this Figure, telomerase activity was present in all lanes except for Lane 2.

Figure 11B is a 12 percent non-denaturing gel of telomerase activity from the immunoprecipitates of the HeLa cell lysates. Prior to testing the immunoprecipitates for telomerase activity, each immunoprecipitate was divided into two aliquots, and RNase was added to one aliquot under conditions described above for RNase incubation. The "+" and "-" symbols at the top of each lane refer to the presence (plus) or absence (minus) of RNase treatment prior to the telomerase assay. Lanes 1 and 2 show telomerase activity in immunoprecipitates with control (non-specific) anti-Myc antibody; Lanes 3 and 4 show a second control antibody (anti-GST); Lanes 5 and 6 show the TP2 specific antibody; Lanes 7-10 show telomerase activity in immunoprecipitates with the anti-TP2 antibody, where the antibody was pre-incubated with 30  $\mu$ g of non-specific peptide 3 (Lanes 7 and 8) or 60  $\mu$ g of non-specific peptide 3 (Lanes 9 and 10); Lanes 11-14 show telomerase activity in immunoprecipitates with the anti-TP2 antibody, where the antibody was pre-incubated with 30  $\mu$ g of TP2 peptide (Lanes 11 and 12) or 60  $\mu$ g TP2 peptide (Lanes 13 and 14). As can be seen, those assays in which RNase was present had a lower amount of telomerase activity as compared with the corresponding assay conducted without RNase, suggesting that telomerase RNA is a necessary component for telomerase activity. Further, those assays in which TP2 was not precipitated with TP2 antibody (*i.e.*, Lanes 1-4 and 11-14) showed a decreased amount of telomerase activity, suggesting that TP2 is associated with telomerase activity.

Figure 11C is a Western blot of the immunoprecipitates. The labels on each Lane are consistent with the labels for Figs. 11A and 11B. The



Western blot was probed with anti-TP2 antiserum in order to detect the presence of TP2 in the immunoprecipitate. Lanes 1 and 2 demonstrate that the non-specific antibodies anti-Myc and anti-GST do not recognize (and thus do not immunoprecipitate) TP2 protein; Lane 3 demonstrates that anti-TP2 antiserum recognizes and immunoprecipitates TP2 protein; Lanes 4 and 5 show that TP2 antiserum which is pre-incubated with non-specific peptide (30  $\mu$ g in Lane 3 and 60  $\mu$ g in Lane 4) is still able to recognize and immunoprecipitate TP2 protein; Lanes 6 and 7 show that TP2 antiserum which is pre-incubated with TP2 specific peptide is not able to immunoprecipitate TP2 protein in the cell lysate.

#### 15 B. Catalytic Activity of TP2

Human embryonic kidney 293 cells (American Type Culture Collection) were cultured in 160 mm plates until they were 50-80 percent confluent. The cells were transfected with either no plasmid (referred to throughout as "MOCK"), wild type plasmid (native, full-length TP2 cDNA referred to throughout as "WT"), or a plasmid containing a single or double point mutation of wild-type TP2 cDNA. Plasmid carrying single or double point mutations were also used for transfection, and are referred to as follows: 5-1 (868D to 868A); 5-2 (869D to 869A); 5-1,2 and (868D to 868A and 869D to 869A).

The DNA constructs used to transfect these cells with TP2 were prepared as described above.

About 24 hours prior to transfection, about 7 X 10<sup>(5)</sup> cells were seeded on to each 100 mm culture dish in about 15 ml of Dulbecco's modified Eagle's medium containing about 10 percent (v/v) fetal calf serum.

For the plasmid transfection, the cells were incubated in about 6 ml of Optimem I reduced serum medium (Gibco-BRL, Grand Island, NY), about 60  $\mu$ l of

Lipofectamine (Invitrogen, San Diego, CA) and about 29  $\mu$ g of TP2 plasmid DNA. After about 4 hours, the medium was replaced with fresh Dulbecco's modified Eagle's medium containing about 10 percent (v/v) fetal calf serum and the cells were harvested about 24 hours later. Cell lysates were prepared essentially according to the method of *Kim et al., supra*. Briefly, the cells were collected by pelleting, and the pellet of cells was washed in PBS buffer and resuspended in a CHAPS detergent buffer. After about 30 minutes on ice, the mixture was spun at about 14,000 X g for 30 minutes, and the supernatants were collected and stored at -80°C.

Telomerase activity assays were conducted as described above for HeLa cells using 1-5  $\mu$ g of cell lysate protein. About 100  $\mu$ g of each 293 cell lysate was then incubated with anti-FLAG M2 affinity gel (mouse IgG1 covalently attached to agarose; Kodak, Rochester, NY). After incubation, the samples were centrifuged to pellet the antibody complex, and washed in hypotonic buffer (Prowse *et al., supra*). The immunoprecipitates were first incubated in the presence or absence of RNase, and the telomerase activity assay was then conducted on the immunoprecipitates as described above.

A 12 percent acrylamide gel with the results of the telomerase assay is shown in Figure 12A. The "+" and "-" signs in Lanes 6-19 of this Figure indicate the presence or absence of RNase, respectively prior to the telomerase assay. In this Figure, "Mock" refers to cells transfected without plasmid; "WT" refers to wild type TP2; and the mutants are labeled according to the description above for their preparation. Lanes 1-5 show lysates from cells transfected with the indicated TP2 gene or control plasmid and these Lanes demonstrate that telomerase activity was present in all cell lysates.

Lanes 6-19 are telomerase assays of TP2 antibody immunoprecipitates of cell lysates. Here, only

the wild type without RNase and the wild type without RNase in which the anti-TP2 antibody was pre-incubated with non-specific peptide (Lanes 8 and 12, respectively) had reasonable levels of telomerase activity. Very little telomerase activity was apparent for the mutant TP2 immunoprecipitates (Lanes 15-19).

Both the immunoprecipitates and the lysates were subjected to Western blot analysis. To determine whether the immunoprecipitates contained TP2, the immunoprecipitate remaining in each aliquot after the activity assay was subjected to SDS-PAGE and then transferred to a PVDF membrane. For detection of TP2 on the Western blot, the blot was first incubated in TBST buffer (Tris-buffered saline in 0.5 percent Tween-20 and 5 percent dried milk) for about 1 hour at room temperature, and then incubated with 0.2 µg/ml of anti-TP2 antibody, followed by horseradish peroxidase conjugated secondary antibody (Amersham, Arlington Heights, IL). The blot was visualized using an ECL kit (Amersham, Arlington Heights, IL).

The results of the Western blot analysis are shown in Figure 12B. The approximate molecular weight of FLAG-TP2 protein is 130 kDa under reducing conditions. FLAG-TP2 protein was detectable in all lysates (Lanes 2-5) except for the MOCK transfected cell lysates (Lane 1), indicating that the 293 cells do not have detectable levels of endogenous TP2 in crude cell lysates, and that both wild type and mutant TP2 proteins are expressed at comparable levels in cell lysates of 293 cells.

All of the immunoprecipitates ("anti-FLAG pellet" lanes in the Figure) except those in Lanes 7 and 9 contained FLAG-TP2 protein as indicated by the approximately 130 kDa band on Lanes 6, 8, and 10-13 of the Western blot. This shows that the TP2 antibody recognizes wild type and mutant TP2 proteins.

The Mock (Lane 7) and the wild type TP2 (Lane 9) in which the anti-FLAG antibody was pre-incubated with about 10 µg of FLAG peptide, did not show detectable TP2 protein. For Lane 9, preparation of FLAG peptide was by standard peptide synthesis methods; the peptide has the sequence set forth in SEQ ID NO: 22; incubation of antiserum with this peptide was for about 30 minutes at room temperature prior to immunoprecipitation.

### C. TRIP1 and TP2 Association

To determine whether TRIP1 and TP2 are associated in the active telomerase complex, Myc-TRIP1, FLAG-TP2, or both were transfected into mouse neuroblastoma N2A cells (American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD, USA). The cells were grown to near confluence in DMEM plus 10 percent fetal calf serum in 100 mm plates, and then transfected with about 10 µg of TP2 cDNA (either MycTRIP1 or FLAGTP2 cDNA) using the Superfect reagent (Qiagen, Chatsworth, CA) according to the manufacturer's directions. Control cells were either not transfected at all ("N2A Lysate" in Figure 13A) or were transfected with plasmid that did not contain TRIP1 or TP2 cDNA ("N2A Mock" in Figure 13A).

About thirty six hours after transfection, cell lysates were prepared as described for HeLa cells using an amount of cell lysate equivalent to about 300 µg of total protein, and telomerase activity assays were conducted as described above for HeLa cells. All cell lysates were immunoprecipitated using the anti-FLAG antibody which recognizes the FLAG epitope of TP2, under conditions described above for anti-FLAG immunoprecipitations. In some cases, the antibody was pre-incubated with anti-FLAG peptide (as described above) or with a Myc peptide (referred to as "non-

specific" peptide). Immunoprecipitates were then subjected to SDS-PAGE, using about one third of the immunoprecipitate per well. After electrophoresis, the proteins were transferred to an Immobilon P membrane (Millipore, Bedford, MA) to generate a Western blot, and the membrane was incubated in PBS plus 0.1 percent Tween-20 (to make "PBST"), and 5 percent non-fat dry milk powder (w:v) as a blocking agent. The membrane was then incubated in a solution of PBST plus 5 percent milk powder and one  $\mu\text{g/ml}$  of anti-Myc or anti-FLAG antibody. Incubation was at room temperature for about 2 hours. Following this step, the membrane was then washed three times for about 10 minutes each time in PBST. Following this step, the membrane was incubated in PBST plus 5 percent milk powder and a rabbit anti-mouse antibody which was conjugated to horse radish peroxidase (Amersham, Arlington Heights, IL). The anti-mouse antibody was added at a 1:1000 dilution. The membrane was then washed again 3 times for 10 minutes each in PBST, and the presence of the anti-Myc or anti-FLAG antibody was detected using the ECL antibody detection kit (Amersham, Arlington Heights, IL) according to the manufacturer's directions.

The results of the Western blot analyses are shown in Figure 13A. In this Figure, the labels at the top of each lane indicate the plasmid used for transfection of the cells- "N2A Mock" (Lane 1) refers to plasmid without TRIP1 or TP2 insert; "MYCTRIPI" (Lane 2) refers to plasmid containing the TRIP1 gene with a MYC label; "FLAGTP2" (Lanes 3-5) refers to full length TP2 containing the FLAG label, where no peptide (Lane 3), specific peptide (Lane 4) or non-specific peptide (Lane 5) was added to the antiserum prior to immunoprecipitation; "MycTRIP1/FLAGTP2" (Lanes 6-8) refers to co-transfection of both the labeled TRIP1 and TP2 genes, where no peptide (Lane 6), specific peptide

(Lane 7) or non-specific peptide (Lane 8) was added to the antiserum prior to immunoprecipitation. The notation "anti-FLAG" and "anti-MYC" on the left indicate which antibody was used to probe each Western blot; the notation on the right shows molecular weight positions, with "TP2" representing the location to which TP2 migrates on the gel/blot, and "TRIP1" representing the location to which TRIP1 migrates on the gel/blot. The results indicate that TRIP1 and TP2 interact, since Lane 6 (minus peptide) and Lane 8 (plus non-specific peptide) show a band that corresponds to TP2 on the anti-FLAG probed Western blot, and a band that corresponds to TRIP1 on the anti-Myc probed Western blot. Therefore, antiserum against TP2 (the anti-FLAG antiserum) immunoprecipitates both TP2 and TRIP1, suggesting that these two components of telomerase are likely associated in the N2A cells.

In addition to Western blots, each immunoprecipitate was tested for telomerase activity using the assay described above, where the immunoprecipitate was first treated with "+" or without "-" RNase under conditions described above.

The results of this assay are shown in Figure 13B. The lanes are labeled as described above. As can be seen, telomerase activity is reduced in those samples exposed to RNase. In addition, Lanes 1-2 show that untransfected N2A cell lysates have endogenous telomerase activity.

Lanes 3-18 are immunoprecipitates using an anti-FLAG antibody (which recognizes the FLAG tag of the recombinant TP2 protein). Lanes 3-4 show that immunoprecipitates of cells transfected with vector only have relatively low levels of telomerase activity.

Lanes 5-6 show that immunoprecipitates of cells transfected with MYC tagged TRIP1 also have relatively low levels of telomerase activity. Lanes 7-

12 are immunoprecipitates of cells transfected with FLAG tagged TP2.

Lane 7 shows increased levels of telomerase activity as compared with Lanes 3 and 5, suggesting that recombiantly expressed TP2 is associated with telomerase activity.

Lanes 9-10 show that telomerase activity associated with recombiantly expressed TP2 is reduced when the anti-FLAG antibody is pre-incubated with a specific FLAG peptide, which verifies that the anti-FLAG antibody specifically immunoprecipitates FLAGTP2 associated telomerase activity. Lanes 11-12 (in which the anti-FLAG antibody was preincubated with non-specific peptide) are control lanes for Lanes 9-10.

Lanes 13-18 are FLAGTP2 antibody immunoprecipitates of cells transfected with both TRIP1 and TP2 labeled DNAs. Lanes 13 shows increased levels of telomerase activity as compared with Lanes 3 and 5, suggesting that recombiantly expressed TP2 is associated with telomerase activity, and that the presence of MycTRIP1 protein in this immunoprecipitate does not affect telomerase activity.

Lanes 15-16 show that telomerase activity associated with recombiantly expressed TP2 is reduced when the anti-FLAG antibody is pre-incubated with a specific FLAG peptide. Lanes 17-18 (in which the anti-FLAG antibody was pre-incubated with non-specific peptide) are control lanes for Lanes 15 and 16.

## 8. In Vitro Reconstitution of Telomerase Activity

As a means of evaluating whether TP2, when combined with telomerase RNA, is biologically active, *in vitro* reconstitution of these components, followed by telomerase activity assays was conducted.

For the *in vitro* assays, human telomerase RNA was obtained as follows. Total human genomic DNA was

prepared from HeLa cells using standard methods. An approximately 520 base pair genomic DNA telomerase DNA fragment was obtained from this genomic DNA by PCR using the following primers:

5

CCCGGGTGGCGGAGGGTGGGC (SEQ ID NO:36)

CGACTTTGGAGGTGCCTTCA (SEQ ID NO:37)

10 Thirty-five rounds of PCR were then conducted in an approximately 50 µl reaction volume using 2 U of Taq polymerase and buffer (Boehringer Mannheim) under the following conditions: 30 seconds at 94°C; 30 seconds at 55°C, and 1.0 minute at 72°C. The PCR product was then  
15 purified from an agarose gel, and this DNA fragment was used as a PCR template to prepare two DNA constructs, each containing the T7 promoter. One construct contained DNA which would generate sense strand human telomerase RNA in a transcription reaction. The other  
20 construct contained DNA which would generate antisense strand human telomerase RNA in a transcription reaction. Both DNA constructs were designed to yield an approximately 450 base pair fragment. The primers used for each PCR were:

25

For Sense RNA:

GGGAAGCTTTAATACGACTCACTATAGGGTGGGCCTGGGAG (SEQ ID NO: 38)

30

CCCGGGGGTTCACAAGCCCCC (SEQ ID NO: 39)

For Antisense RNA:

35 GGGAAGCTTTAATACGACTCACTATAGGGGGTTCACAAGCCCCC (SEQ ID NO: 40)



CCCGGGTGGGCCTGGGAG

(SEQ ID NO: 41)

These constructs were then used to prepare telomerase RNA by transcription using the DNA constructs as templates. About 5-15  $\mu$ g of template DNA, 40 mM Tris-HCl, pH 8.0, 6 mM MgCl<sub>2</sub>, 10 mM DTT, 1 mM of each ribonucleotide, 2 mM spermidine, 350 U of RNA Guard (Boehringer Mannheim), and 500 U of T7 RNA polymerase (Boehringer Mannheim) were incubated together in a final volume of about 200  $\mu$ L for about 1.5 hours at 37°C. After incubation, the samples were extracted with one volume of phenol:chloroform (1:1), chloroform, ethanol precipitated, and resuspended in sterile water.

After the transcription reaction, the products were resolved on a denaturing acrylamide gel, and the full length telomerase RNAs (approximately 450 base pairs) was identified by UV shadowing, excised, and purified by elution of the RNA from crushed acrylamide slice using an Acrodisc (Gelman Sciences, Ann Arbor, MI).

Separately, full length human TP2 cDNA was inserted into the vector pCR3 (Stratagene, La Jolla, CA) using standard cloning and ligation methods (see Example 5 above; the construct used had the FLAG tag at both the 5' and 3' ends).

Approximately 0.5  $\mu$ L of TP2pCR3 construct was placed in about 50  $\mu$ g of *in vitro* rabbit reticulocyte translation buffer ("TNT" *in vitro* reconstitution kit; Promega, Madison, WI). This mixture was then divided into 10  $\mu$ L aliquots, and various amounts of each DNA construct for use in generating human telomerase RNA was then added. The reaction was started by adding 1  $\mu$ L T7 polymerase (according to the manufacturer's instructions), and was allowed to proceed for about 1.5 hours at about 30°C. The reaction was stopped by placing each sample on ice. Approximately one  $\mu$ L of each

sample was then assayed for telomerase activity using the telomerase assay described above. Various RNAs other than human sense RNA were used as controls for the *in vitro* translations. These included antisense human telomerase RNA (prepared as described above; results shown in Lanes 10-13 of Figure 14), sense mouse telomerase RNA (prepared by transcription from a vector containing the full length mouse telomerase RNA preceded by the T7 promoter (see Harrington *et al.*, *Science* 275:973--977 [1997])); results shown in Lanes 18-21 of Figure 14), and transfer RNA (Sigma Chemical Co., St. Louis, MO; results shown in Lanes 14-17 of Figure 14).

Hela cell lysate (prepared as described in Example 7 above) was used as a positive control for the telomerase assay the amount of lysate used per telomerase assay was equivalent to approximately 5  $\mu$ g of protein). The HeLa cell lysate was assayed in the absence (Figure 14; Lane 1) or presence (Figure 14; Lane 2) of RNase. Negative controls included 1  $\mu$ L of TNT lysate with no TP2 DNA construct added (Figure 14; Lane 3); 1  $\mu$ L of TNT lysate with TP2 DNA but no added RNA (Figure 14; Lane 4); 1  $\mu$ L of TNT lysate with 0.001, 0.005, 0.01, or 0.1  $\mu$ g of human telomerase RNA (Lanes 22-25 of Figure 14).

Experimental *in vitro* reconstitution samples contained TP2 DNA template in about 10  $\mu$ L of TNT reaction mixture in the presence of 0.01, 0.05, 0.1, or 1  $\mu$ g of either sense human telomerase RNA (Lanes 5-8 of Figure 14) of which 1  $\mu$ L was used for the telomerase assay. Lane 9 of Figure 14 shows the same reaction as Lane 8, except that about 1  $\mu$ g RNase was added prior to the telomerase assay.

As can be seen in Figure 14, only those samples containing both telomerase sense RNA and TP2 protein in which no RNase was present (Lanes 5-8) had telomerase activity.

The ability of the TP2 mutants to abolish telomerase activity in the reticulocyte translation system was evaluated in experiments comparable to those described immediately above, but using a DNA construct encoding either full length wild type TP2, or one of the TP2 mutants 5-1, 5-2, or 5-1,2 (prepared as described in Example 5) in each *in vitro* translation reaction. Positive and negative controls were used as described above. The results are shown in Figure 15. Lanes 1-9 are the same as for Figure 15. Lanes 10-13 show TP2 mutant 5-1; Lanes 14-17 show TP2 mutant 5-2; and Lanes 18-21 show TP2 mutant 5-1.2. For Lanes 10-21, the amount of telomerase RNA or control RNA present in the assay is indicated.

The results of this series of assays indicate that wild type TP2 and sense telomerase RNA are required for telomerase activity (Lanes 5-8).

To evaluate whether TP2 clone 32 (see Example 5; this TP2 clone is less than full length) had catalytic telomerase activity *in vitro*, an *in vitro* reconstitution assay was performed using the TNT kit (Promega, Madison, WI) essentially as described above. For the reconstitution assays, about 50  $\mu$ L of TNT reticulocyte extract was mixed with about 1  $\mu$ g of sense strand human telomerase RNA, either 1  $\mu$ g of full length TP2 cDNA, or 1  $\mu$ g of clone 32 TP2 cDNA in the pCR3 vector ("short TP2"; also referred to as "TPs") and about 1  $\mu$ L of T7 RNA polymerase. The reaction was conducted for about 1.5 hours at about 30°C, and was stopped by placing the samples on ice. About 1  $\mu$ L of each sample was then assayed for telomerase activity using the telomerase assay described in Example 7.

The results of the telomerase assay are shown in Figure 17A. Lanes 1 and 2 show HeLa cell lysate (about 5  $\mu$ g protein) incubated without (Lane 1) or with (Lane 2) 1  $\mu$ g of RNase. Lanes 3 and 4 show the results

of full length TP2 plus human telomerase RNA (Lane 3) and short TP2 plus human telomerase RNA (Lane 4). Lanes 5-8 show the results of reconstitution and telomerase assays performed in the absence of either telomerase RNA (Lanes 5 and 6), in the absence of TP2 (Lane 7), or in the absence of RNA and TP2 (Lane 8). As can be seen, both full length TP2 and clone 32 TP2 (TPs) were active in the telomerase assay in the presence of telomerase RNA.

The Western blot shown in Figure 17B shows that TP2 protein was generated in the reticulocyte mixture when full length TP2 cDNA or clone 32 TP2 cDNA ("TPs") was present.

The effect of TRIP1 on telomerase activity was evaluated as follows. An *in vitro* reconstitution assay using the TNT lysate system was performed using full length TP2 cDNA plus human telomerase RNA as described above. Separately, an *in vitro* reconstitution assay using the TNT lysate system and full length TRIP1 cDNA was performed as described above. After incubation of each assay in the presence of T7 RNA polymerase, about 6  $\mu$ L of the TP2 containing extract was added to the same volume of TRIP1 extract, and the mixture was incubated on ice for about 30 minutes. As a control, about 6  $\mu$ L of TP2 extract was added to an extract that had neither telomerase RNA nor TRIP1 or TP2 cDNA (or protein), and this mixture was incubated on ice as well.. Either 1 or 2  $\mu$ L of each mixture, as well as the TP2 only extract, was assayed for telomerase activity using the methods described above.

The results are shown in Figure 18. "no DNA" refers to reticulocyte extract to which neither DNA encoding telomerase RNA, nor DNA encoding TP2 or TRIP1 was added; "-TP1" refers to extract containing telomerase RNA and TP2 protein only; and "+TP1" refers to the mixture of extracts containing telomerase RNA,

TRIP1 protein, and TP2 protein. As can be seen, those extracts containing TP2, TRIP1, and telomerase RNA (Lanes 5 and 6) had enhanced telomerase activity as compared to the extracts with TP2 and telomerase RNA only (Lanes 3 and 4).

#### 9. Dominant-Negative Assays of TP2

Each of the four TP2 expression cassettes were excised from the vector pCR3 and were inserted into the pIRES-EGFP expression vector (Clontech, Palo Alto, CA) by digesting each of the pCR3/TP2 vectors with *KpnI* and *XbaI* and inserting the resultant TP2 cDNAs into the pIRES-EGFP vector previously modified to introduce a *NheI* site (compatible with *KpnI*) and *XbaI* site between the inherent *NotI* and *EcoRI* sites of the vector. The pIRES-EGFP vector is designed such that the expression of the test gene (in this case full length wild type or mutant TP2) is directly related to the amount of Green Fluorescent Protein ("GFP") produced by the expression cassette of the vector.

Human embryonic kidney 293 cells (see Example 8B) were grown in 100 mm plates in DMEM with 10% FCS to near confluence, and were then transfected with 11 ug of a TP2/pIRES-EGFP vector per plate, where the TP2 in the vector was either wild type, or one of the three mutants. In addition, control cells were transfected with pIRES-EGFP vector that did not contain a TP2 insert. Transfection was accomplished using Superfect reagent (Qiagen, Chatsworth, CA) according to the manufacturer's directions. Forty eight hours later, the transfected cells were removed from the plates by treating with trypsin for five minutes. The cells were resuspended in PBS supplemented with 2 percent fetal calf serum. Next, each of the transfected cell

populations was separately sorted using standard Fluorescence Activated Cell Sorting (FACS) techniques.

FACS sorting permitted the separation of each of the 5 populations of transfected cells according the level of GFP expressed in each cell. Each transfected cell containing a construct that expressed low levels of GFP was pooled into a population termed "low". Cells with higher levels of GFP were pooled into a population termed "high". In addition, a population that did not contain transfected DNA was identified, and is referred to as "non".

Each of the 15 populations of cells from the five transfected vectors were then pelleted by centrifugation at 15,000 x G and lysed in 100  $\mu$ L of telomerase lysis buffer.

Figure 16A shows a Western blot of lysates from each of the 15 cell populations. The Western blot was prepared as described above in Example 7, and was probed with the anti-FLAG antibody (Kodak, Rochester, NY). Lanes 1-3 show lysates from "non", "low" and "high" GFP expressing cell populations, respectively, of the 5.1 TP2 mutants. As can be seen, there is a direct relationship between the expression levels of GFP (based on FACS analysis, see above) and expression levels of the TP2 protein. The "non" (Lane 1) does not have detectable FLAG 5.1 TP2 mutant protein, while "low" (Lane 2) shows a low level of FLAG 5.1 TP2 mutant protein, and "high" (Lane 3) shows a higher level of FLAG 5.1 TP2 mutant protein. Lanes 4-6 show a similar relationship with FLAG 5.2 TP2 mutant protein, and these lanes contain "non", "low", and "high" lysates, respectively. Lanes 7-9 show a similar relationship with FLAG 5.1/5.2 TP2 mutant protein, while Lanes 10-12 show a similar relationship with FLAG wild type TP2 protein. Lanes 13-15 show that cells transfected with

pIRES-EGFP without TP2 insert do not express a detectable level of TP2 protein.

Figure 16B shows the associated telomerase activity for each of the 15 cell population lysates described for Figure 16A. Lanes 1-3 show an inverse relationship between the expression of the TP2 mutant protein (Figure 16A) and telomerase activity. The same inverse relationship is seen for all of the mutants (Lanes 4-6 for mutant 5.2, and Lanes 7-9 for double mutant 5-1,2). In contrast, telomerase activity levels in the lysates from the wild type TP2 expressors are essentially the same, regardless of the amount of wild type TP2 present in the lysates. Similarly, telomerase activity is essentially constant in the cells transfected with IRES-EGFP vector alone (Lanes 13-15). This result suggests that biologically inactive TP2 mutants, when expressed in a cell that normally has telomerase activity, can be used to decrease or suppress such telomerase activity.

#### Deposit of TRIP1 cDNA

*E. coli* cells containing the plasmid pCR3 with the insert TRIP1MycTag3 (encoding mouse full length TRIP1 polypeptide) has been deposited with the ATCC (American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD, USA) on November 8, 1996 as accession number 98250. In addition, four separate clones of *E. coli* cells containing the plasmid pSPORT into which a portion of the human TRIP1 cDNA coding sequence were deposited with the ATCC on the same date. Clone 15 contains cDNA encoding amino acids 1046-2627 and has ATCC accession number 98254; clone 54 contains cDNA encoding amino acids 423-1467 and has ATCC accession number 98253; clone 63 contains cDNA encoding amino acids 1346-2488 and has ATCC accession number 98252; and

A-433A

- 97 -

clone 96 contains cDNA encoding amino acids 1-567 and has ATCC accession number 98251.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100



We Claim:

1. A TP2 nucleic acid molecule encoding a  
5 polypeptide selected from the group consisting of:
  - (a) the nucleic acid molecule of SEQ ID NO:13;
  - (b) the nucleic acid molecule that is  
nucleotides 1920-2820 of SEQ ID NO:13;
  - (c) the nucleic acid molecule of SEQ ID NO:19
  - 10 (d) a nucleic acid molecule encoding the  
polypeptide of SEQ ID NO:14, or a biologically active  
fragment thereof;
  - (e) a nucleic acid molecule encoding the  
polypeptide of SEQ ID NO:20, or a biologically active
  - 15 fragment thereof;
  - (f) a nucleic acid molecule that encodes a  
polypeptide that is at least 90 percent identical to the  
polypeptide of SEQ ID NO:14;
  - (g) a nucleic acid molecule that encodes a
  - 20 polypeptide that is at least 90 percent identical to the  
polypeptide of SEQ ID NO:20;
  - (h) a nucleic acid molecule that hybridizes  
under stringent conditions to any of (a)-(g) above; and
  - (i) a nucleic acid molecule that is the
  - 25 complement of any of (a)-(g) above.
2. The nucleic acid molecule that is SEQ ID  
NO:13 or SEQ ID NO:19.
- 30 3. The nucleic acid molecule that is  
nucleotides 1920-2820 of SEQ ID NO:13.
4. A nucleic acid molecule encoding the  
polypeptide of SEQ ID NO:14 of SEQ ID NO:20.

35

5. A nucleic acid molecule selected from the group consisting of: nucleotides 1-1689 of SEQ ID NO:13, nucleotides 1-1920 of SEQ ID NO:13, nucleotides 1920-2820 of SEQ ID NO:13, nucleotides 2089-2820 of SEQ ID NO:13, and nucleotides 2089-2859 of SEQ ID NO:13.

6. A nucleic acid molecule encoding amino acids 640-940 of the polypeptide of SEQ ID NO:14.

7. A vector comprising the nucleic acid molecule of claim 1.

8. A vector comprising the nucleic acid molecule of claim 2.

9. A vector comprising the nucleic acid molecule of claim 3.

10. A vector comprising the nucleic acid molecule of claim 4.

11. A vector comprising the nucleic acid molecule of claim 5.

12. A vector comprising the nucleic acid molecule of claim 6.

13. A host cell comprising the vector of claim 7.

14. A host cell comprising the vector of claim 8.

15. A host cell comprising the vector of claim 9.

16. A host cell comprising the vector of claim 10.

17. A host cell comprising the vector of claim 11.

18. A host cell comprising the vector of claim 12.

19. A process for producing a TP2 polypeptide comprising the steps of:

- (a) expressing a polypeptide encoded by the nucleic acid of claim 1 in a suitable host; and
- (b) isolating the polypeptide.

20. The process of claim 19 wherein the polypeptide is SEQ ID NO:14 or SEQ ID NO:20.

21. The process of claim 19 wherein the polypeptide is amino acids 640-940 of SEQ ID NO:14.

22. A TP2 polypeptide selected from the group consisting of:

- (a) the polypeptide of SEQ ID NO:14;
- (b) the polypeptide that is amino acids 640-940 of SEQ ID NO:14;
- (c) the polypeptide of SEQ ID NO:20; and
- (d) a polypeptide that is at least 90 percent identical to any of the polypeptides of (a)-(c).

23. A TP2 polypeptide that is the polypeptide of SEQ ID NO:14, SEQ ID NO:20, or a biologically active fragment thereof.

24. A TP2 polypeptide selected from the group consisting of: amino acids 1-563 of SEQ ID NO:14; amino

acids 1-640 of SEQ ID NO:14; amino acids 640-940 of SEQ ID NO:14; amino acids 696-940 of SEQ ID NO:14; and amino acids 696-953 of SEQ ID NO:14.

5                   25. The TP2 polypeptide of claim 22 that does not possess an amino terminal methionine.

26. A method of increasing proliferation of a cell, comprising expressing a nucleic acid encoding TP2  
10 or a biologically active fragment thereof, in the cell.

27. A method of increasing telomerase activity in a cell, comprising expressing a TP2 gene, or a biologically active fragment thereof, in the cell.  
15

28. A method of decreasing telomerase in a cell, comprising expressing a TP2 mutant in a cell, wherein the mutant does not have TP2 biological activity.  
20

29. A nucleic acid molecule encoding a mutant TP2 polypeptide, wherein the codon for aspartic acid at amino acid position 868 or 869 is changed to a codon for alanine.  
25

30. A nucleic acid molecule encoding a mutant TP2 polypeptide, wherein the codons for aspartic acid at amino acid positions 868 and 869 are changed to codons for alanine.  
30

31. A polypeptide encoded by the nucleic acid molecule of claim 29.

32. A polypeptide encoded by the nucleic acid molecule of claim 30.  
35

NOVEL GENES ENCODING TELOMERASE PROTEINS  
ABSTRACT OF THE DISCLOSURE

Disclosed are nucleic acid molecules encoding  
5 polypeptides of the telomerase complex. Also disclosed  
are methods of preparing the nucleic acid molecules and  
polypeptides, and methods of using these molecules.

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70  
75  
80  
85  
90  
95  
100  
105  
110  
115  
120  
125  
130  
135  
140  
145  
150  
155  
160  
165  
170  
175  
180  
185  
190  
195  
200  
205  
210  
215  
220  
225  
230  
235  
240  
245  
250  
255  
260  
265  
270  
275  
280  
285  
290  
295  
300  
305  
310  
315  
320  
325  
330  
335  
340  
345  
350  
355  
360  
365  
370  
375  
380  
385  
390  
395  
400  
405  
410  
415  
420  
425  
430  
435  
440  
445  
450  
455  
460  
465  
470  
475  
480  
485  
490  
495  
500  
505  
510  
515  
520  
525  
530  
535  
540  
545  
550  
555  
560  
565  
570  
575  
580  
585  
590  
595  
600  
605  
610  
615  
620  
625  
630  
635  
640  
645  
650  
655  
660  
665  
670  
675  
680  
685  
690  
695  
700  
705  
710  
715  
720  
725  
730  
735  
740  
745  
750  
755  
760  
765  
770  
775  
780  
785  
790  
795  
800  
805  
810  
815  
820  
825  
830  
835  
840  
845  
850  
855  
860  
865  
870  
875  
880  
885  
890  
895  
900  
905  
910  
915  
920  
925  
930  
935  
940  
945  
950  
955  
960  
965  
970  
975  
980  
985  
990  
995

**FIGURE 1A**

ATGGAAAACTCCATGGGCATGTGTCTGCCCATCCAGACATCCTCTCCT  
TGGAGAACCGGTGCCTGGCTATGCTCCCTGACTTACAGCCCTTGGAGAA  
ACTACATCAGCATGTATCTACCCACTCAGATATCCTCTCCTTGAAGAAC  
CAGTGCCTAGCCACGCTTCCTGACCTGAAGACCATGGAAAAACCACATG  
GATATGTGTCTGCCACCCAGACATCCTCTCCTTGGAGAACCAGTGCCT  
GGCCACACTTTCTGACCTGAAGACCATGGAGAAACCACATGGACATGTT  
TCTGCCACCCAGACATCCTCTCCTTGGAGAACCGGTGCCTGGCCACCC  
TCCCTAGTCTAAAGAGCACTGTGTCTGCCAGCCCCCTTGTTCCAGAGTCT  
ACAGATATCTCACATGACGCAAGCTGATTTGTACCGTGTGAACAACAGC  
AATTGCCTGCTCTCTGAGCCTCCAAGTTGGAGGGCTCAGCATTTCTCTA  
AGGGACTAGACCTTTCAACCTGCCCTATAGCCCTGAAATCCATCTCTGC  
CACAGAGACAGCTCAGGAAGCAACTTTGGGTCGTTGGTTTGATTCAGAA  
GAGAAGAAAGGGGCAGAGACCCAAATGCCTTCTTATAGTCTGAGCTTGG  
GAGAGGAGGAGGAGGTGGAGGATCTGGCCGTGAAGCTCACCTCTGGAGA  
CTCTGAATCTCATCCAGAGCCTACTGACCATGTCCTTCAGGAAAAGAAG  
ATGGCTCTACTGAGCTTGCTGTGCTCTACTCTGGTCTCAGAAGTAAACA  
TGAACAATACATCTGACCCACCCCTGGCTGCCATTTTTGAAATCTGTCTG  
TGAACCTGCCCTCCTGGAGCCTGAGTTTATCCTCAAGGCATCTTTGTAT  
GCCAGGCAGCAGCTGAACGTCCGGAATGTGGCCAATAACATCTTGGCCA

**FIGURE 1B**

TTGCTGCTTTCTTGCCGGCGTGTCGCCCCACCTGCGACGATATTTCTG  
TGCCATTGTCCAGCTGCCTTCTGACTGGATCCAGGTGGCTGAGCTTTAC  
CAGAGCCTGGCTGAGGGAGATAAGAATAAGCTGGTGCCCCTGCCCGCCT  
GTCTCCGTACTGCCATGACGGACAAATTTGCCCAGTTTGACGAGTACCA  
GCTGGCTAAGTACAACCCTCGGAAGCACCGGGCCAAGAGACACCCCCGC  
CGGCCACCCCGCTCTCCAGGGATGGAGCCTCCATTTTCTCACAGATGTT  
TTCCAAGGTACATAGGGTTTCTCAGAGAAGAGCAGAGAAAGTTTGAGAA  
GGCCGGTGATACAGTGTCAGAGAAAAAGAATCCTCCAAGGTTACCCCTG  
AAGAAGCTGGTTCAGCGACTGCACATCCACAAGCCTGCCCAGCACGTTC  
AAGCCCTGCTGGGTTACAGATACCCCTCCAACCTACAGCTCTTTTCTCG  
AAGTCGCCTTCCTGGGCCTTGGGATTCTAGCAGAGCTGGGAAGAGGATG  
AAGCTGTCTAGGCCAGAGACCTGGGAGCGGGAGCTGAGCCTACGGGGGA  
ACAAAGCGTCGGTCTGGGAGGAACTCATTGAAAATGGGAAGCTTCCCTT  
CATGGCCATGCTTCGGAACCTGTGCAACCTGCTGCGGGTTGGAATCAGT  
TCCCGCCACCATGAGCTCATTCTCCAGAGACTCCAGCATGGGAAGTCGG  
TGATCCACAGTCGGCAGTTTCCATTCAGATTTCTTAACGCCCATGATGC  
CATTGATGCCCTCGAGGCTCAACTCAGAAATCAAGCATTGCCCTTTCCT  
TCGAATATAACACTGATGAGGCGGATACTAACTAGAAATGAAAAGAACC  
GTCCCAGGCGGAGGTTTCTTTGCCACCTAAGCCGTCAGCAGCTTCGTAT

**FIGURE 1C**

GGCAATGAGGATACCTGTGTTGTATGAGCAGCTCAAGAGGGAGAAGCTG  
AGAGTACACAAGGCCAGACAGTGGAAATATGATGGTGAGATGCTGAACA  
GGTACCGACAGGCCCTAGAGACAGCTGTGAACCTCTCTGTGAAGCACAG  
CCTGCCCCCTGCTGCCAGGCCGCACTGTCTTGGTCTATCTGACAGATGCT  
AATGCAGACAGGCTCTGTCCAAAGAGCAACCCACAAGGGCCCCCGCTGA  
ACTATGCACTGCTGTTGATTGGGATGATGATCACGAGGGCGGAGCAGGT  
GGACGTCGTGCTGTGTGGAGGTGACACTCTGAAGACTGCAGTGCTTAAG  
GCAGAAGAAGGCATCCTGAAGACTGCCATCAAGCTCCAGGCTCAAGTCC  
AGGAGTTTGATGAAAATGATGGATGGTCCCTGAATACTTTTGGGAAATA  
CCTGCTGTCTCTGGCTGGCCAAAGGGTTCCTGTGGACAGGGTCATCCTC  
CTTGGCCAAAGCATGGATGATGGAATGATAAATGTGGCCAAACAGCTTT  
ACTGGCAGCGTGTGAATTCCAAGTGCCTCTTTGTTGGTATCCTCCTAAG  
AAGGGTACAATACCTGTCAACAGATTTGAATCCCAATGATGTGACACTC  
TCAGGCTGTACTGATGCGATACTGAAGTTCATTGCAGAGCATGGGGCCT  
CCCATCTTCTGGAACATGTGGGCCAAATGGACAAAATATTCAAGATTCC  
ACCACCCCCAGGAAAGACAGGGGTCCAGTCTCTCCGGCCACTGGAAGAG  
GACACTCCAAGCCCCCTTGGCTCCTGTTTCCCAGCAAGGATGGCGCAGCA  
TCCGGCTTTTCATTTTCATCCACTTTCCGAGACATGCACGGGGAGCGGGA  
CCTGCTGCTGAGGTCTGTGCTGCCAGCACTGCAGGCCCGAGCGGCCCT



**FIGURE 1D**

CACCGTATCAGCCTTCACGGAATCGACCTCCGCTGGGGCGTCACTGAGG  
 AGGAGACCCGTAGGAACAGACAACCTGGAAGTGTGCCTTGGGGAGGTGGA  
 GAACGCACAGCTGTTTGTGGGGATTCTGGGCTCCCGTTATGGATACATT  
 CCCCCAGCTACAACCTTCCTGACCATCCACACTTCCACTGGGCCCAGC  
 AGTACCCCTTCAGGGCGCTCTGTGACAGAGATGGAGGTGATGCAGTTCCT  
 GAACCGGAACCAACGTCTGCAGCCCTCTGCCAAGCTCTCATCTACTTC  
 CGGGATTCCAGCTTCCTCAGCTCTGTGCCAGATGCCTGGAAATCTGACT  
 TTGTTTCTGAGTCTGAAGAGGCCGCATGTCGGATCTCAGAACTGAAGAG  
 CTACCTAAGCAGACAGAAAGGGATAACCTGCCGCAGATACCCCTGTGAG  
 TGGGGGGGTGTGGCAGCTGGCCGGCCCTATGTTGGCGGGCTGGAGGAGT  
 TTGGGCAGTTGGTTCTGCAGGATGTATGGAATATGATCCAGAAGCTCTA  
 CCTGCAGCCTGGGGCCCTGCTGGAGCAGCCAGTGTCCATCCCAGACGAT  
 GACTTGGTCCAGGCCACCTTCCAGCAGCTGCAGAAGCCACCGAGTCCTG  
 CCCGGCCACGCCTTCTTCAGGACACAGTGCAACAGCTGATGCTGCCCCA  
 CGGAAGGCTGAGCCTGGTGACGGGGCAGTCAGGACAGGGCAAGACAGCC  
 TTCCTGGCATCTCTTGTGTCAGCCCTGCAGGCTCCTGATGGGGCCAAGG  
 TGGCACCATTAGTCTTCTTCCACTTTTCTGGGGCTCGTCCTGACCAGGG  
 TCTTGCCCTCACTCTGCTCAGACGCCTCTGTACCTATCTGCGTGGCCAA  
 CTAAAAGAGCCAGGTGCCCTCCCCAGCACCTACCGAAGCCTGGTGTGGG

**FIGURE 1E**

AGCTGCAGCAGAGGCTGCTGCCCAAGTCTGCTGAGTCCCTGCATCCTGG  
CCAGACCCAGGTCCTGATCATCGATGGGGCTGATAGGTTAGTGGACCAG  
AATGGGCAGCTGATTTTCAGACTGGATCCCAAAGAAGCTTCCCCGGTGTG  
TACACCTGGTGCTGAGTGTGTCTAGTGATGCAGGCCTAGGGGAGACCCT  
TGAGCAGAGCCAGGGTGCCACGTGCTGGCCTTGGGGCCTCTGGAGGCC  
TCTGCTCGGGCCCGGCTGGTGAGAGAGGAGCTGGCCCTGTACGGGAAGC  
GGCTGGAGGAGTCACCATTTAACAACCAGATGCGACTGCTGCTGGTGAA  
GCGGGAATCAGGCCGGCCGCTCTACCTGCGCTTGGTCACCGATCACCTG  
AGGCTCTTCACGCTGTATGAGCAGGTGTCTGAGAGACTCCGGACCCTGC  
CTGCCACTGTCCCCCTGCTGCTGCAGCACATCCTGAGCACACTGGAGAA  
GGAGCACGGGCCTGATGTCCTTCCCCAGGCCTTGACTGCCCTAGAAGTC  
ACACGGAGTGGTTTGGACTGTGGACCAGCTGCACGGAGTGCTGAGTGTGT  
GGCGGACACTACCGAAGGGGACTAAGAGCTGGGAAGAAGCAGTGGCTGC  
TGGTAACAGTGGAGACCCCTACCCCATGGGCCCCGTTTGCCTGCCTCGTC  
CAGAGTCTGCGCAGTTTGTCTAGGGGAGGGCCCTCTGGAGCGCCCTGGTG  
CCCGGCTGTGCCTCCCTGATGGGCCCTGAGAACAGCAGCTAAACGTTG  
CTATGGGAAGAGGCCAGGGCTAGAGGACACGGCACACATCCTCATTGCA  
GCTCAGCTCTGGAAGACATGTGACGCTGATGCCTCAGGCACCTTCCGAA  
GTTGCCCTCCTGAGGCTCTGGGAGACCTGCCTTACCACCTGCTCCAGAG

# FIGURE 1F

CGGGAACCGTGGACTTCTTTTGAAGTTCCTTACCAACCTCCATGTGGTG  
GCTGCACACTTGGAATTGGGTCTGGTCTCTCGGCTCTTGGAGGCCCATG  
CCCTCTATGCTTCTTCAGTCCCCAAAGAGGAACAAAAGCTCCCCGAGGC  
TGACGTTGCAGTGTTTCGCACCTTCCTGAGGCAGCAGGCTTCAATCCTC  
AGCCAGTACCCCCGGCTCCTGCCCCAGCAGGCAGCCAACCAGCCCCCTGG  
ACTCACCTCTTTGCCACCAAGCCTCGCTGCTCTCCCGGAGATGGCACCT  
CCAACACACACTACGATGGCTTAATAAACCCCGGACCATGAAAAATCAG  
CAAAGCTCCAGCCTGTCTCTGGCAGTTTCCTCATCCCCTACTGCTGTGG  
CCTTCTCCACCAATGGGCAAAGAGCAGCTGTGGGCACTGCCAATGGGAC  
AGTTTACCTGTTGGACCTGAGAACTTGGCAGGAGGAGAAGTCTGTGGTG  
AGTGGCTGTGATGGAATCTCTGCTTGTTTGTTCTCTCCGATGATACAC  
TCTTTCTTACTGCCTTCGACGGGCTCCTGGAGCTCTGGGACCTGCAGCA  
TGGTTGTCGGGTGCTGCAGACTAAGGCTCACCAGTACCAAATCACTGGC  
TGCTGCCTGAGCCCAGACTGCCGGCTGCTAGCCACCGTGTGCTTGGGAG  
GATGCCTAAAGCTGTGGGACACAGTCCGTGGGCAGCTGGCCTTCCAGCA  
CACCTACCCCAAGTCCCTGAACTGTGTTGCCTTCCACCCAGAGGGGCAG  
GTAATAGCCACAGGCAGCTGGGCTGGCAGCATCAGCTTCTTCCAGGTGG  
ATGGGCTCAAAGTCACCAAGGACCTGGGGGCACCCGGAGCCTCTATCCG  
TACCTTGGCCTTCAATGTGCCTGGGGGGGTGTGGCTGTGGGCCGGCTG

**FIGURE 1G**

GACAGTATGGTGGAGCTGTGGGCCTGGCGAGAAGGGGCACGGCTGGCTG  
CCTTCCCTGCCCACCATGGCTTTGTTGCTGCTGCGCTTTTCCTGCATGC  
GGGTGCCAGTTACTGACGGCTGGAGAGGATGGCAAGGTTTCAGGTGTGG  
TCAGGGTCTCTGGGTGCGCCCCGTGGGCACCTGGGTTCCTTTCTCTCT  
CTCCTGCCCTCTCTGTGGCACTCAGCCCAGATGGTGATCGGGTGGCTGT  
TGGATATCGAGCGGATGGCATTAGGATCTACAAAATCTCTTCAGGTTC  
CAGGGGGCTCAGGGTCAGGCACTGGATGTGGCAGTGTCCGCCCTGGCCT  
GGCTAAGCCCCAAGGTATTGGTGAGTGGTGCAGAAGATGGGTCCTTGCA  
GGGCTGGGCACTCAAGGAATGCTCCCTTCAGTCCCTCTGGCTCCTGTCC  
AGATTCCAGAAGCCTGTGCTAGGACTGGCCACTTCCCAGGAGCTCTTGG  
CTTCTGCCTCAGAGGATTTACAGTGCAGCTGTGGCCAAGGCAGCTGCT  
GACGCGGCCACACAAGGCAGAAGACTTTCCCTGTGGCACTGAGCTGCGG  
GGACATGAGGGCCCTGTGAGCTGCTGTAGTTTCAGCACTGATGGAGGCA  
GCCTGGCCACCGGGGGCCGGGATCGGAGTCTCCTCTGCTGGGACGTGAG  
GACACCCAAAACCCCTGTTTTGATCCACTCCTTCCCTGCCTGTCACCGT  
GACTGGGTCACTGGCTGTGCCTGGACCAAAGATAACCTACTGATATCCT  
GCTCCAGTGATGGCTCTGTGGGGCTCTGGGACCCAGAGTCAGGACAGCG  
GCTTGGTCAGTTCCTGGGTTCATCAGAGTGCTGTGAGCGCTGTGGCAGCT  
GTGGAGGAGCACGTGGTGTCTGTGAGCCGGGATGGGACCTTGAAAGTGT

**FIGURE 1H**

GGGACCATCAAGGCGTGGAGCTGACCAGCATCCCTGCTCACTCAGGACC  
CATTAGCCACTGTGCAGCTGCCATGGAGCCCCGTGCAGCTGGACAGCCT  
GGGTCAGAGCTTCTGGTGGTAACCGTCGGGCTAGATGGGGCCACACGGT  
TATGGCATCCACTCTTGGTGTGCCAAACCCACACCCTCCTGGGACACAG  
CGGCCCAGTCCGTGCTGCTGCTGTTTCAGAAACCTCAGGCCTCATGCTG  
ACCGCCTCTGAGGATGGTTCTGTACGGCTCTGGCAGGTTTCCTAAGGAAG  
CAGATGACACATGTATACCAAGGAGTTCTGCAGCCGTCACTGCTGTGGC  
TTGGGCACCAGATGGTTCCATGGCAGTATCTGGAAATCAAGCTGGGGAA  
CTAATCTTGTGGCAGGAAGCTAAGGCTGTGGCCACAGCACAGGCTCCAG  
GCCACATTGGTGCTCTGATCTGGTCCTCGGCACACACCTTTTTTTGTCCT  
CAGTGCTGATGAGAAAATCAGCGAGTGGCAAGTGAAACTGCGGAAGGGT  
TCGGCACCCGGAAATTTGAGTCTTCACCTGAACCGAATTCTACAGGAGG  
ACTTAGGGGTGCTGACAAGTCTGGATTGGGCTCCTGATGGTCACTTTCT  
CATCTTGGCCAAAGCAGATTTGAAGTTACTTTGCATGAAGCCAGGGGAT  
GCTCCATCTGAAATCTGGAGCAGCTATACAGAAAATCCTATGATATTGT  
CCACCCACAAGGAGTATGGCATATTTGTCCTGCAGCCCAAGGATCCTGG  
AGTTCTTTCTTTCTTGAGGCAAAAGGAATCAGGAGAGTTTGAAGAGAGG  
CTGAACTTTGATATAAACTTAGAGAATCCTAGTAGGACCCTAATATCGA  
TAACTCAAGCCAAACCTGAATCTGAGTCCTCATTTTTTGTGTGCCAGCTC

**FIGURE 1I**

TGATGGGATCCTATGGAACCTGGCCAAATGCAGCCCAGAAGGAGAATGG  
ACCACAGGTAACATGTGGCAGAAAAAAGCAAACACTCCAGAAACCCAAA  
CTCCAGGGACAGACCCATCTACCTGCAGGGAATCTGATGCCAGCATGGA  
TAGTGATGCCAGCATGGATAGTGAGCCAACACCACATCTAAAGACACGG  
CAGCGTAGAAAGATTCACTCGGGCTCTGTTCACAGCCCTCCATGTGCTAC  
CTGAGTTGCTGGTGACAGCTTCGAAGGACAGAGATGTTAAGCTATGGGA  
GAGACCCAGTATGCAGCTGCTGGGCCTGTTCCGATGCGAAGGGTCAGTG  
AGCTGCCTTGGAACCTTGGCTGGGCGCTAACTCCACCCTGCAGCTTGCCG  
TGGGAGACGTGCAGGGCAATGTGTACTTTCTGAATTGGGAA

**FIGURE 2A**

ATGGAGAAGCTCTGTGGGCATGTGCCTGGCCATTCAGACATCCTCTCCT  
TGAAGAACCGGTGCCTGACCATGCTCCCTGACCTCCAGCCCCTGGAGAA  
AATACATGGACATAGATCTGTCCACTCAGACATCCTTTCCTTGGAGAAC  
CAGTGTCTGACCATGCTCTCTGACCTCCAGCCCACGGAGAGAATAGATG  
GGCATATATCTGTCCACCCAGACATCCTCTCCTTGGAGAATCGGTGCCT  
GACCATGCTCCCTGACCTCCAGCCTCTGGAGAAGCTATGTGGACATATG  
TCTAGTCATCCAGACGTCCTTTCTTTGGAAAACCAATGTCTAGCTACTC  
TCCCCACTGTAAAGAGCACTGCATTGACCAGCCCCCTTGCTCCAGGGTCT  
TCACATATCTCATAACGGCACAAGCTGATCTGCATAGCCTGAAAACCTAGC  
AACTGCCTGCTCCCTGAGCTTCCTACCAAGAAGACTCCATGTTTCTCTG  
AGGAACTAGACCTTCCACCTGGACCCAGGGCCCTGAAATCCATGTCTGC  
TACAGCTCAAGTCCAGGAAGTAGCCTTGGGTCAATGGTGTGTCTCCAAA  
GAAAAGGAATTTCAAGAAGAAGAAAGCACAGAAGTCCCRATGCCTTTGT  
ACAGTCTAAGCTTGGAAGAAGAAGAAGTGGAGGCACCGGTCTTAAAACT  
CACATCTGGAGACTCTGGCTTTCATCCTGAAACCACTGACCAGGTCCTT  
CAGGAGAAGAAGATGGCTCTCTTGACCTTACTCTGCTCTGCTCTGGCCT  
CAAATGTGAATGTGAAAGATGCATCTGACCTTACCCGGGCATCCATCCT  
TGAAGTCTGTAGTGCCCTGGCCTCCTTGGAACCGGAGTTCATCCTTAAG  
GCATCTTTGTATGCTCGGCAGCAACTTAACCTCCGGGACATCGCCAATA

**FIGURE 2B**

CAGTTCTGGCTGTGGCTGCCCTCTTGCCAGCCTGCCGCCCCCATGTACG  
ACGGTATTACTCCGCCATTGTTACCTGCCTTCAGACTGGATCCAGGTA  
GCCGAGTTCTACCAGAGCCTGGCAGAAGGGGATGAGAAGAAGTTGGTGT  
CCCTGCCTGCCTGTCTCCGAGCTGCCATGACCGACAAATTTGCCGAGTT  
TGATGAGTACCAGCTAGCTAAGTACAACCCACGGAAACATCGGTCCAAG  
AGGCGGTCCCGCCAGCCACCCCGCCCTCAAAGACAGAACGTCCATTTT  
CAGAGAGAGGGAAATGTTTTCCAAAGAGCCTTTGGCCCCCTTAAAAATGA  
ACAGATTACGTTTGAAGCAGCTTATAATGCAATGCCAGAGAAAAACAGG  
CTACCACGGTTCACCTCTGAAGAAGTTGGTAGAGTATCTACATATCCACA  
AGCCTGCTCAGCACGTCCAGGCCCTGCTGGGCTACAGGTACCCAGCCAC  
CCTAGAGCTCTTTTCTCGGAGTCACCTCCCTGGGCCGTGGGAGTCTAGC  
AGAGCTGGTCAGCGGATGAAGCTCCGAAGGCCAGAGACCTGGGAGCGGG  
AGCTGAGTTTACGGGGAAACAAAGCTTCTGTGTGGGAGGAGCTCATAGA  
CAATGGGAAACTGCCCTTCATGGCCATGCTCCGGAACCTGTGTAACCTG  
CTGCGGACTGGGATCAGTGCCCGCCACCATGAACTCGTTCTCCAGAGAC  
TCCAGCATGAGAAATCTGTGGTTCACAGTCGGCAGTTTCCATTTCAGATT  
CCTTAATGCTCATGACTCTATCGATAAACTTGAGGCTCAGCTCAGAAGC  
AAAGCATCACCTTCCCTTCCAATACAACATTGATGAAACGGATAATGA  
TTAGAAACTCAAAAAAAAAATAGGAGGCCTGCCAGTCGGAAGCACCTGTG



# FIGURE 2C

CACCCTGACGCGCCGGCAGCTTCGGGCAGCAATGACTATACCTGTGATG  
TATGAGCAGCTCAAGCGGGAGAACTGAGGCTGCACAAGGCCAGACAAT  
GGAAGTGTGATGTTGAGTTGCTGGAGCGCTATCGCCAGGCCCTGGAAAC  
AGCTGTGAACCTCTCAGTAAAGCACAACTATCCCCGATGCCTGGCCGA  
ACCCTCTTGGTCTATCTCACAGATGCAAATGCCGACAGGCTCTGTCCCA  
AGAGTCACTCACAAGGGCCTCCCCTGAACTATGTGCTGCTGCTGATCGG  
AATGATGGTGGCTCGAGCCGAGCAAGTGACTGTTTGCTTGTGTGGGGGA  
GGATTTGTGAAGACACCGGTACTTACAGCCGATGAAGGCATCCTGAAGA  
CTGCCATCAAACCTTCAGGCTCAAGTCCAGGAGTTAGAAGGCAATGATGA  
GTGGCCCCCTGGACACTTTTGGGAAGTATCTGCTGTCTCTGGCTGTCCAA  
AGGACCCCCATTGACAGGGTCATCCTGTTTGGTCAAAGGATGGATACCG  
AGCTCCTGAAAGTAGCCAAACAGATTATCTGGCAGCATGTGAATTCCAA  
GTGCCTCTTTGTTGGTGTCTCCTACAGAAAACACAGTACATATCACCA  
AATTTGAATCCCAACGATGTGACGCTCTCAGGCTGCACTGACGGGATCC  
TGAAATTCATTGCCGAACATGGAGCCTCTCGTCTCCTGGAACATGTGGG  
ACAACTAGATAAACTATTCAAGATCCCCCACCACCAGGAAAGACACAG  
GCACCGTCTCTCCGGCCGCTGGAGGAGAACATCCCTGGTCCCTTGGGTC  
CTATTTCCCAGCATGGATGGCGCAATATCCGGCTTTTCATTTTCATCCAC  
TTTCCGTGACATGCATGGGGAGCGAGATTTGCTGATGAGATCTGTTCTG

# FIGURE 2D

CCCGCACTGCAGGCCAGAGTGTTCCTCCACCGCATCAGTCTTCACGCCA  
 TTGACCTGCGCTGGGGTATCACAGAGGAAGAGACCCGCAGGAACAGACA  
 ACTGGAAGTGTGCCTTGGGGAGGTGGAGAACTCACAGCTGTTTCGTGGGG  
 ATTCTGGGCTCCCGCTATGGCTACATTCCCCCAGCTATGATCTTCCTG  
 ATCATCCCCACTTTCACTGGACCCATGAGTACCCTTCAGGGCGATCCGT  
 GACAGAGATGGAGGTGATGCAATTCCTGAACCGTGGCCAACGCTCGCAG  
 CCTTCGGCCCAAGCTCTCATCTACTTCCGAGATCCTGATTTCTTTAGCT  
 CTGTGCCAGATGCCTGGAAACCTGACTTTATATCTGAGTCAGAAGAAGC  
 TGCACATCGGGTCTCAGAGCTGAAGAGATATCTACACGAACAGAAAGAG  
 GTTACCTGTGCGCAGCTACTCCTGTGAATGGGGAGGTGTAGCGGCTGGCC  
 GGCCCTATACTGGGGGCCTGGAGGAGTTTGGACAGTTGGTTCTCCAGGA  
 TGTGTGGAGCATGATCCAGAAGCAGCACCTGCAGCCTGGGGCCCAGTTG  
 GAGCAGCCAACATCCATCTCAGAAGACGATTTGATCCAGACCAGCTTTC  
 AGCAGCTGAAGACCCCAACGAGTCCGGCACGGCCACGCCTTCTTCAGGA  
 TACAGTGCAGCAGCTGTTGCTGCCCCATGGGAGGCTGAGCCTAGTGACT  
 GGGCAGGCAGGACAGGGAAAGACTGCCTTTCTGGCATCCCTTGTGTCTG  
 CCCTGAAGGTCCCTGACCAGCCCAATGAGCCCCCGTTTCGTTTTCTTCCA  
 CTTTGCAGCAGCCCGCCCTGACCAGTGTCTTGCTCTCAACCTCCTCAGA  
 CGCCTCTGTACCCATCTGCGTCAAAAACCTGGGAGAGCTGAGTGCCCTCC

# FIGURE 2E

CCAGCACTTACAGAGGCCTGGTGTGGGAACTGCAGCAGAAGTTGCTCCT  
CAAATTCGCTCAGTCGCTGCAGCCTGCTCAGACTTTGGTCCTTATCATC  
GATGGGGCAGATAAGTTGGTGGATCGTAATGGGCAGCTGATTTCAGACT  
GGATCCCCAAGTCTCTTCCGCGGCGAGTACACCTGGTGCTGAGTGTGTC  
CAGTGACTCAGGCCTGGGTGAGACCCCTTCAGCAAAGTCAGGGTGCTTAT  
GTGGTGGCCTTGGGCTCTTTGGTCCCATCTTCAAGGGCTCAGCTTGTGA  
GAGAAGAGCTAGCACTGTATGGGAAACGACTGGAGGAGTCACCTTTTAA  
CAACCAGATGCGGCTGCTGCTGGCAAAGCAGGGTTCAAGCCTGCCATTG  
TACCTGCACCTTGTCACCTGACTACCTGAGGCTCTTCACACTGTATGAAC  
AGGTGTCTGAGAGACTTCGAACCCTGCCCCGCCACTCTCCCACTGCTCTT  
GCAGCACATCCTGAGCACCTTGGAGCAAGAACATGGCCATGATGTCCTT  
CCTCAGGCTTTGACTGCCCTTGAGGTCACACGAAGTGGTCTGACTGTGG  
ACCAGCTACATGCAATCCTGAGCACATGGCTGATCTTGCCCAAGGAGAC  
TAAGAGCTGGGAAGAAGTGCTGGCTGCCAGTCACAGTGGAACCCCTTTC  
CCCTTGTGTCCATTTGCCTACCTTGTCCAGAGTCTACGCAGTTTACTAG  
GGGAGGGCCCAGTGGAGCGCCCTGGTGCCCGTCTCTGCCTCTCTGATGG  
GCCCCTGAGGACAACAATTAAACGTCGCTATGGGAAAAGGCTGGGGCTA  
GAGAAGACTGCGCATGTCCTCATTGCAGCTCACCTCTGGAAGACGTGTG  
ATCCTGATGCCTCGGGCACCTTCCGAAGTTGCCCTCCTGAGGCTCTGAA

# FIGURE 2F

AGATTTACCTTACCACCTGCTCCAGAGCGGGAACCATGGTCTCCTTGCC  
GAGTTTCTTACCAATCTCCATGTGGTTGCTGCATATCTGGAAGTGGGTC  
TAGTCCCCGACCTCTTGGAGGCTCATGTGCTCTATGCTTCTTCAAAGCC  
TGAAGCCAACCAGAAGCTCCCAGCGGCAGATGTTGCTGTTTTCCATACC  
TTCCTGAGACAACAGGCTTCACTCCTTACCCAGTATCCTTTGCTCCTGC  
TCCAGCAGGCAGCTAGCCAGCCTGAAGAGTCACCTGTTTGCTGCCAGGC  
CCCCCTGCTCACCCAGCGATGGCACGACCAGTTCACACTGAAATGGATT  
AATAAACCCCAGACCCTGAAGGGTCAGCAAAGCTTGTCTCTGACAATGT  
CCTCATCCCCAACTGCTGTGGCCTTCTCCCCGAATGGGCAAAGAGCAGC  
TGTGGGGACCGCCAGTGGGACAATTTACCTGTTGAACTTGAAAACCTGG  
CAGGAGGAGAAGGCTGTGGTGAGTGGCTGTGACGGGATTTCTCTTTTG  
CATTCCTTTTCGGACACTGCCCTTTTCCTTACTACCTTCGACGGGCACCT  
AGAGCTTTGGGACCTGCAACATGGTTGTTGGGTGTTTCAGACCAAGGCC  
CACCAGTACCAAATCACTGGCTGCTGCCTGAGCCCAGACCGCCGCCTGC  
TGGCCACTGTGTGTTTGGGAGGATACCTAAAGCTGTGGGACACAGTCCG  
AGGACAGCTGGCTTTTCAGTACACCCATCCAAAGTCTCTCAACTGCGTT  
GCCTTCCACCCAGAGGGGCAGGTGGTAGCCACAGGCAGCTGGGCTGGCA  
GCATTACCTTCTTCCAGGCAGATGGACTCAAAGTCACCAAGGAAGTAGG  
GGCCCCCGGACCCTCTGTCTGTAGTTTGGCATTCAACAAACCTGGGAAG

# FIGURE 2G

ATTGTGGCTGTGGGCCGGATAGATGGGACAGTGGAGCTGTGGGCCTGGC  
AAGAGGGTGCCCGGCTGGCGGCCTTCCCTGCACAGTGTGGCTGTGTCTC  
TGCTGTTCTTTTCTTGCATGCTGGAGACCGGTTCCCTGACTGCTGGAGAA  
GATGGCAAGGCTCAGTTATGGTCAGGATTTCTTGGCCGGCCCAGGGGT  
GCCTGGGCTCTCTTCCCTCTTTCTCCTGCACTCTCGGTGGCTCTCAACCC  
AGACGGTGACCAGGTGGCTGTTGGGTACCGAGAAGATGGCATTAACATC  
TACAAGATTTCTTCAGGTTCACAGGGGCTCAGCATCAAGAGCTAAATG  
TGGCGGTGTCTGCACTGGTGTGGCTGAGCCCTAGTGTTTTGGTGAGTGG  
TGCAGAAGATGGATCCCTGCATGGTTGGATGTTCAAGGGAGACTCCCTT  
CATTCCCTGTGGCTGTTGTGCGAGATAACAGAAGCCTGTGCTGGGACTGG  
CTGCCTCCCGGGAACATCATGGCTGCTGCCTCAGAGGACTTCACTGTGAG  
ACTGTGGCCCAGACAGCTGCTGACACAGCCACATGTGCATGCGGTAGAG  
TTGCCCTGTTGTGCTGAACTCCGGGGACACGAGGGGCCAGTGTGCTGCT  
GTAGCTTCAGCCCTGATGGAGGCATCTTGGCCACAGCTGGCAGGGATCG  
GAATCTCCTTTGCTGGGACATGAAGATAGCCCAAGCCCCTCTCCTGATT  
CACACTTTCTCGTCCTGTCATCGTGACTGGATCACTGGCTGTGCGTGGA  
CCAAAGACAACATCCTGGTCTCCTGCTCGAGTGATGGCTCTGTGGGACT  
CTGGAACCCAGAGGCAGGGCAGCAACTTGGCCAGTTCTCAGGCCACCAG  
AGTGCCGTGAGCGCCGTGGTTGCTGTGGAGGAACACATTGTATCTGTGA

**FIGURE 2H**

GCCGAGATGGGACCTTGAAAGTGTGGGACCATCAGGGTGTGGAGCTGAC  
CAGCATCCCTGCCCATTCCGGACCCATCAGCCAGTGTGCAGCTGCTCTG  
GAGCCCCGCCAGGGGGACAGCCTGGATCAGAGCTTCTGGTGGTGA CTG  
TTGGACTAGATGGGGCCACAAAGTTGTGGCATCCCCTGTTGGTGTGCCA  
AATACGTACTCTCCAGGGACACAGTGGCCCAGTCACAGCAGCTGCTGCT  
TCAGAGGCCTCAGGCCTCCTGCTGACCTCAGATGATAGCTCTGTACAGC  
TCTGGCAGATAACAAAGGAAGCAGATGATTCATACAAACCTAGGAGTTC  
TGTGGCCATCACTGCTGTGGCATGGGCACCGGATGGTTCTATGGTGGTG  
TCCGGAAATGAAGCCGGGGA ACTGACACTGTGGCAGCAAGCCAAGGCTG  
TGGCTACCGCACAGGCTCCAGGCCGCGTCAGTCACCTGATCTGGTACTC  
GGCAAATTCATTCTTCGTTCTCAGTGCTAATGAAAACGTCAGCGAGTGG  
CAAGTGGGACTGAGGAAAGGTTCAACGTCCACCAGTTCAGTCTTCATC  
TGAAGAGAGTTCTGCAGGAGGACTGGGGAGTCTTGACAGGTCTGGGTCT  
GGCCCCTGATGGCCAGTCTCTCATCTTGATGAAAGAGGATGTGGAATTA  
CTAGAGATGAAGCCTGGGTCTATTCCATCTTCTATCTGCAGGAGGTATG  
GAGTACATTCTTCAATACTGTGCACCAGCAAGGAGTACGGCTTGTTCTA  
CCTGCAGCAGGGGGACTCCGGATTACTTTCTATATTGGAGCAAAGGAG  
TCAGGGGAGTTTGAAGAGATCCTGGACTTCAATCTGAACTTAAATAATC  
CTAATGGGTCCCCAGTATCAATCACTCAGGCCAAACCTGAGTCTGAATC

**FIGURE 2I**

ATCCCTTTTGTGCGCCACCTCTGATGGGATGCTGTGGAACCTTATCTGAA  
TGTACCTCAGAGGGAGAATGGATCGTAGATAACATTTGGCAGAAAAAG  
CAAAAAACCTAAAACTCAGACTCTGGAGACAGAGTTGTCCCCGCACTC  
AGAGTTGGATTTTTCCATTGATTGCTGGATTGATCCCACAAATTTAAAG  
GCACAGCAGTGTA AAAAGATCCACTTGGGCTCTGTCACAGCCCTCCATG  
TGCTTCCGGGATTGCTGGTGACAGCTTCGAAGGACAGAGATGTTAAGCT  
GTGGGAGAGACCCAGTATGCAGCTGCTGGGCTTGTTCCGATGTGAAGGG  
CCAGTGAGCTGTCTGGAACCTTGGATGGAGCCCAGCTCTCCCCTGCAGC  
TTGCTGTGGGAGACACACAAGGAACTTGTATTTTCTATCTTGGGAA

# FIGURE 3A

MEKLHGHVSAHPDILSLENRCLAMLPDLQPLEKLHQHVSTHSDILSLKN  
 QCLATLPDLKTMEKPHGYVSAHPDILSLENQCLATLSDLKTMEKPHGHV  
 SAHPDILSLENRCLATLPSLKSTVSASPLFQSLQISHMTQADLYRVNNS  
 NCLLSEPPSWRAQHFSKGLDLSTCPIALKSISATETAQEATLGRWFDSE  
 EKKGAETQMPYSYSLSGEEEEVEDLAVKLTSGDSESHPEPTDHVLQEEK  
 MALLSLLCSTLVSEVNMNNTSDPTLAAIFEICRELALLEPEFILKASLY  
 ARQQLNVRNVANNILAIAAFLPACRPHLRRYFCAIVQLPSDWIQVAELY  
 QSLAEGDKNKLVPPLPACLRRTAMTDKFAQFDEYQLAKYNPRKHKRAKRHPR  
 RPPRSPGMEPPFSHRCFPRYIGFLREEQRKFEEKAGDTVSEKKNPPRFTL  
 KKLVSRLHIIHKPAQHVQALLGYRYPSNLQLFSRSRLPGPWDSSRAGKRM  
 KLSRPETWERELSLRGNKASVWEELIENGKLPFMAMLRNLCNLLRVGIS  
 SRHHELILQRLQHGKSVIHSRQFPFRFLNAHDAIDALEAQLRNQALPFP  
 SNITLMRRILTRNEKNRPRRRFLCHLSRQQLRMAMRIPVLYEQLKREKL  
 RVHKARQWKYDGEMLNRYRQALETAVNLSVKHSLPLLPGRTVLVYLTDA  
 NADRLCPKSNPQGPPLNYALLIGMMITRAEQVDVVLCCGGDTLKTAVLK  
 AEEGILKTAIKLQAQVQEFDENDGWSLNTFGKYLLSLAGQRVPVDRVIL  
 LGQSMDDGMINVAKQLYWQRVNSKCLFVGILLRRVQYLSTDLPNDVTL  
 SGCTDAILKFIAEHGASHLLEHVGQMDKIFKIPPPPGKTGVQSLRPLEE  
 DTPSPLAPVSQQGWSIRLFISS'TFRDMHGERDLLLLRSVLPALQARAAP



# FIGURE 3B

HRISLHGIDLRWGVTEEEETRRNRQLEVCLGEVENAQLFVGILGSRYGYI  
 PPSYNLPDHPHFHWAQQYPSGRSVTEMEVMQFLNRNQRLQPSAQALIYF  
 RDSSFLSSVPDAWKSDFVSESEEAAXRISELKSYLSRQKGITCRRYPCE  
 WGGVAAGRPHYVGGLEEFQQLVLQDVWNMIQKLYLQPGALLEQPVSI PDD  
 DLVQATFQQQLQKPPSPARPRLLQDTVQXLMLPHGRLSLVTGQSGQGKTA  
 FLASLVSALQAPDGAKVAXLVFFHFSGARPDQGLALTLLRRLCTYL RGQ  
 LKEPGALPSTYRSLVWELQQRLLPKSAESLHPGQTQVLIIDGADRLVDQ  
 NGQLISDWIPKKLPRCVHLVLSVSSDAGLGETLEQSQGAHVLALGPLEA  
 SARARLVREELALYGKRLEESPFNNQMRLLLVKRESGRPLYLRLVTDHL  
 RLFTLYEQVSERLRTL PATVPLLLQHILSTLEKEHGPDVLPQALTALEV  
 TRSGLTVDQLHGVL SVWRTL PKGTKSWEEAVAAGNSGDPYPMGPFACLV  
 QSLRSLLGEGPLERPGARLCLPDGPLRTAAKRCYGKRPGLEDTAHILIA  
 AQLWKTCDADASGTFRSCPPEALGDL PYHLLQSGNRGLLSKFLTNLHV  
 AAHLELGLVSRLL EAHALYASSVPKEEQKLPEADVAVFRTFLRQQASIL  
 SQYPRLLPQQAANQPLDSPLCHQASLLSRRWHLQHTLRWLNKPRTMKNQ  
 QSSSLSLAVSSSPTAVAFSTNGQRAAVGTANGTVYLLDLRTWQEEKSVV  
 SGCDGISACLFLSDDTLFLTAFDGLLELWDLQHGCRLVLTQKAHQYQITG  
 CCLSPDCRLLATVCLGGCLKLWDTVRGQLAFQHTYPKSLNCVAFHPEGQ  
 VIATGSWAGSISFFQVDGLKVTKDLGAPGASIRTLAFNVPGGVVAVGRL

# FIGURE 3C

DSMVELWAWREGARLAAFPAHHGFVAAALFLHAGCQLLTAGEDGKVQVW  
 SGSLGRPRGHLGSLSLSPALSVALSPDGDRVAVGYRADGIRIYKISSGS  
 QGAQGQALDVAVSALAWLSPKVLVSGAEDGSLQGWALKECSLQSLWLLS  
 RFQKPVLGLATSQELLASASEDFTVQLWPRQLLTRPHKAEDFPCGTEL  
 R  
 GHEGPVSCCSFSTDGGSLATGGRDRSLLCWDVTRTPKTPVLIHSFPACHR  
 DWVTGCAWTKDNLLISCSSDGSVGLWDPESGQRLGQFLGHQSAVSAVAA  
 VEEHVVSVRDGTKVWDHQGVELTSIPAHS GPISHCAAAMEPRAAGQP  
 GSELLVVTVGLDGATRLWHPLLCVQTHLTLLGHSGPVRAAAVSETSGLML  
 TASEDGSVRLWQVPKEADDTCTIPRSSAAVTAVAWAPDGSMASGNQAGE  
 LILWQEAKAVATAQAPGHIGALIWSSAHTFFVLSADEKISEWQVKLRKG  
 SAPGNLSLHLNRILQEDLGVLTSLDWAPDGHFLILAKADLKLKCMKPGD  
 APSEIWSSYTENPMILSTHKEYGIFVLQPKDPGVLSFLRQKESGEFEER  
 LNF DINLENPSRTLISITQAKPESESSFLCASSDGILWNLAKCSPEGEW  
 TTGNMWQKKANTPETQTPGTDPSTCRESDASMDSDASMDSEPTPHLKTR  
 QRRKIHSGSVTALHVLPELLVTASKDRDVKLWERPSMQLLGLFRCEGSV  
 SCLEPWLGANSTLQLAVGDVQGNVYFLNWE

# FIGURE 4A

MEKLCGHVPGHSDILSLKNRCLTMLPDLQPLEKIHGHRSVHSDILSLEN  
 QCLTMLSDLQPTERIDGHISVHPDILSLENRCLTMLPDLQPLEKLCGHM  
 SSHPDVLSLENQCLATLPTVKSTALTSPLLQGLHISHTAQADLHSLKTS  
 NCLLPELPTKKTPCFSEELDLPPGPRALKSMSATAQVQEVALGQWCVSK  
 EKEFQEEESTEVPMPLYSLSEEEEEVEAPVLKLTSGDSGFHPETTDQVL  
 QEKKMALLTLLCSALASNVNVKDASDLTRASILEVCSALASLEPEFILK  
 ASLYARQQNLNRDIANTVLAVAALLPACRPHVRRYSAIVHLPSDWIQV  
 AEFYQSLAEGDEKKLVSLPACLRAAMTDKFAEFDEYQLAKYNPRKHRSK  
 RRSRQPPRPQKTERPFSERGKCFPKSLWPLKNEQITFEAAYNAMPEKNR  
 LPRFTLKKLV EYLHIHKPAQH VQALLGYRYPATLELFSRSHLPGPWESS  
 RAGQRMKLRRPETWERELSLRGNKASVWEELIDNGKLPFMAMLRNLCNL  
 LRTGISARHHELVLQRLQHEKSVVHSRQFPFRFLNAHDSIDKLEAQLRS  
 KASPFPSNTTLMKRIMIRNSKKNRRPASRKHLCTLTRRQLRAAMTIPVM  
 YEQLKREKLRLHKARQWNCDVELLERYRQALETAVNLSVKHNLSMPMPGR  
 TLLVYLTDANADRLCPKSHSQGPPLNYVLLLLIGMMVARAEQVTVCLCGG  
 GFVKTPVLTAD EGILKTAIKLQAQVQELEGNDEWPLDTFGKYLLSLAVQ  
 RTPIDRVILFGQRM DTELLKVAKQIIWQH VNSKCLFVG VLLQKTQYISP  
 NLNPNDVTLSGCTDGILKFIAEHGASRLLEHVGQLDKLFKIPPPPGKTQ  
 APSLRPLEENIPGPLGPISQHGWRNIRLFISSTFRDMHGERDLLMRSVL

## FIGURE 4B

PALQARVFPHRISLHAIDLRWGITTEEETRRNRQLEVCLGEVENSQLFVG  
ILGSRYGYIPPSYDLDPHHPHFHWTHEYPSGRSVTEMEVMQFLNRGQRSQ  
PSAQALIYFRDPDFLSSVPDAWKPDFISESEEAHRVSELKRYLHEQKE  
VTCRSYSCEWGGVAAGRPYTGGLLEFGQLVLQDVWSMIQKQHLQPGAQL  
EQPTSISEDLLIQTSFQQLKTPTSPARPRLLQDTVQQLLLPHGRLSLVT  
GQAGQGKTAFLASLVSALKVPDQPNPPFVFFHFAAARPDQCLALNLLR  
RLCTHLRQKLGELSALPSTYRGLVWELQQKLLKFAQSLQPAQTLVLI  
DGADKLVDNRNGQLISDWIPKSLPRRVHLVLSVSSDSGLGETLQQSQGAY  
VVALGSLVPSSRAQLVREELALYGKRLEESPFNNQMRLLLAKQGSSLPL  
YLHLVTDYLRRLFTLYEQVSERLRTLPLPLLLQHILSTLEQEHGHDVL  
PQALTALEVTRSGLTVDQLHAILSTWLILPKETKSWEVLAASHSGNPF  
PLCPFAYLVQSLRSLLGEGPVERPGARLCLSDGPLRTTIKRRYGKRLGL  
EKTAHVLIAAHLWKTCDPDASGTFRSCPPEALKDLPYHLLQSGNHGLLA  
EFLTNLHVVAAYLEVGLVPDLLEAHVLYASSKPEANQKLPAADVAVFHT  
FLRQQASLLTQYPLLLLQQAASQPEESPVCCQAPLLTQRWHDQFTLKI  
NKPQTLKGQQLSLTMSSSPTAVAFSPNGQRAAVGTASGTIYLLNLKTW  
QEEKAVVSGCDGISSFAFLSDTALFLTTFDGHLELWDLQHGCWVFQTKA  
HQYQITGCCLSPDRLLATVCLGGYLKLWDTVRGQLAFQYTHPKSLNCV  
AFHPEGQVVATGSWAGSITFFQADGLKVTKELGAPGPSVCSLAFNKP GK

# FIGURE 4C

IVAVGRIDGTVELWAWQEGARLAAFPAQCGCVSAVLFLHAGDRFLTAGE  
DGKAQLWSGFLGRPRGCLGSLPLSPALSVALNPDGDQVAVGYREDGINI  
YKISSGSQGPQHQELNVAVSALVWLSPSVLVSGAEDGSLHGWMFKGDSL  
HSLWLLSRYQKPVLGLAASRELMAAASEDFTVRLWPRQLLTQPHVHAVE  
LPCCAELRGHEGPVCCCSFSPDGGILATAGRDRNLLCWDMKIAQAPLLI  
HTFSSCHRDWITGCAWTKDNILVSCSSDGSVGLWNPEAGQQLGQFSGHQ  
SAVSAVVAVEEHIVSVSRDGTCLKVWDHQVELTSSIPAHSGPISQCAAAL  
EPRPGGQPGSELLVVTVGLDGATKLWHPLLVCQIRTLQGHSGPVTAAAA  
SEASGLLLTSDDSSVQLWQIPKEADDSYKPRSSVAITAVAWAPDGSMVV  
SGNEAGELTLWQQAQAVATAQAPGRVSHLIWYSANSFFVLSANENVSEW  
QVGLRKGSTSTSSSLHLKRVLQEDWGVLTGLGLAPDGQSLILMKEDVEL  
LEMKPGSIPSSICRRYGVHSSILCTSKEYGLFYLQQGDSGLLSILEQKE  
SGEFEEILDFNLNLNNPNGSPVSITQAKPESESSLLCATSDGMLWNLSE  
CTSEGEWIVDNIWQKKAKKPKTQTLETELSPHSELDFSIDCWIDPTNLK  
AQQCKKIHLGSVTALHVLPGLLVTASKDRDVKLWERPSMQLLGLFRCEG  
PVSCLEPWMEPSSPLQLAVGDTQGNLYFLSWE

**FIGURE 5A**

CACGCGTCCGGGCAGCGCTGCGTCCTGCTGCGCACGTGGGAAGCCCTGG  
CCCCGGCCACCCCCGCGATGCCGCGCGCTCCCCGCTGCCGAGCCGTGCG  
CTCCCTGCTGCGCAGCCACTACCGCGAGGTGCTGCCGCTGGCCACGTTC  
GTGCGGCGCCTGGGGCCCCAGGGCTGGCGGCTGGTGCAGCGCGGGGACC  
CGGCGGCTTTCCGCGCGCTGGTGGCCCAGTGCCCTGGTGTGCGTGCCCTG  
GGACGCACGGCCGCCCCCGCCGCCCCCTCCTTCCGCCAGGTGTCCTGC  
CTGAAGGAGCTGGTGGCCCGAGTGCTGCAGAGGCTGTGCGAGCGCGGCG  
CGAAGAACGTGCTGGCCTTCGGCTTCGCGCTGCTGGACGGGGCCCCGCGG  
GGGCCCCCCCCGAGGCCTTCACCACCAGCGTGCGCAGCTACCTGCCCAAC  
ACGGTGACCGACGCACTGCGGGGGAGCGGGGCGTGGGGGCTGCTGCTGC  
GCCGCGTGGGCGACGACGTGCTGGTTCACCTGCTGGCACGCTGCGCGCT  
CTTTGTGCTGGTGGCTCCCAGCTGCGCCTACCAGGTGTGCGGGCCGCCG  
CTGTACCAGCTCGGCGCTGCCACTCAGGCCCGGCCCCCGCCACACGCTA  
GTGGACCCCGAAGGCGTCTGGGATGCGAACGGGCCTGGAACCATAGCGT  
CAGGGAGGCCGGGGTCCCCCTGGGCCTGCCAGCCCCGGGTGCGAGGAGG  
CGCGGGGGCAGTGCCAGCCGAAGTCTGCCGTTGCCCAAGAGGCCAGGC  
GTGGCGCTGCCCCCTGAGCCGGAGCGGACGCCCGTTGGGCAGGGGTCTTG  
GGCCCACCCGGGCAGGACGCGTGACCGAGTGACCGTGGTTTCTGTGTG  
GTGTCACCTGCCAGACCCGCCGAAGAAGCCACCTCTTTGGAGGGTGCGC

**FIGURE 5B**

TCTCTGGCACGCGCCACTCCCACCCATCCGTGGGCGCCAGCACACGC  
GGGCCCCCATCCACATCGCGGCCACCACGTCCCTGGGACACGCCTTGT  
CCCCCGGTGTACGCCGAGACCAAGCACTTCCTCTACTCCTCAGGCGACA  
AGGAGCAGCTGCGGCCCTCCTTCCTACTCAGCTCTCTGAGGCCCAGCCT  
GACTGGCGCTCGGAGGCTCGTGGAGACCATCTTTCTGGGTTCAGGCC  
TGGATGCCAGGGACTCCCCGCAGGTTGCCCCGCCTGCCCCAGCGCTACT  
GGCAAATGCGGCCCTGTTCCTGGAGCTGCTTGGAACACGCGCAGTG  
CCCCTACGGGGTGCTCCTCAAGACGCACTGCCCGCTGCGAGCTGCGGTC  
ACCCCAGCAGCCGGTGTCTGTGCCCGGGAGAAGCCCCAGGGCTCTGTGG  
CGGCCCCCGAGGAGGAGGACACAGACCCCCGTGCTTGGTGAGCTGCT  
CCGCCAGCACAGCAGCCCCTGGCAGGTGTACGGCTTCGTGCGGGCCTGC  
CTGCGCCGGCTGGTGCCCCCAGGCCTCTGGGGCTCCAGGCACAACGAAC  
GCCGCTTCCTCAGGAACACCAAGAAGTTCATCTCCCTGGGGAAGCATGC  
CAAGCTCTCGCTGCAGGAGCTGACGTGGAAGATGAGCGTGCGGGACTGC  
GCTTGGCTGCGCAGGAGCCCAGGGGTTGGCTGTGTTCCGGCCGCAGAGC  
ACCGTCTGCGTGAGGAGATCCTGGCCAAGTTCCTGCACTGGCTGATGAG  
TGTGTACGTCGTCGAGCTGCTCAGGTCTTTCTTTTATGTCACGGAGACC  
ACGTTTCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCA  
AGTTGCAAAGCATTGGAATCAGACAGCACTTGAAGAGGGTGAGCTGCG

**FIGURE 5C**

GGAGCTGTCGGAAGCAGAGGTCAGGCAGCATCGGGAAGCCAGGCCCCGCC  
CTGCTGACGTCCAGACTCCGCTTCATCCCCAAGCCTGACGGGCTGCGGC  
CGATTGTGAACATGGACTACGTTCGTGGGAGCCAGAACGTTCCGCAGAGA  
AAAGAGGGCCGAGCGTCTCACCTCGAGGGTGAAGGCACTGTTTCAGCGTG  
CTCAACTACGAGCGGGCGCGGGCGCCCCGGCCTCCTGGGCGCCTCTGTGC  
TGGGCCTGGACGATATCCACAGGGCCTGGCGCACCTTCGTGCTGCGTGT  
GCGGGCCCAGGACCCGCCGCTGAGCTGTACTTTGTCAAGGTGGATGTG  
ACGGGCGCGTACGACACCATCCCCAGGACAGGCTCACGGAGGTCATCG  
CCAGCATCATCAAACCCAGAACACGTACTGCGTGCGTCGGTATGCCGT  
GGTCCAGAAGGCCGCCCATGGGCACGTCCGCAAGGCCTTCAAGAGCCAC  
GTCTCTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTTCGTGGCTC  
ACCTGCAGGAGACCAGCCCGCTGAGGGATGCCGTCGTCATCGAGCAGAG  
CTCCTCCCTGAATGAGGCCAGCAGTGGCCTCTTCGACGTCTTCCTACGC  
TTCATGTGCCACCACGCCGTGCGCATCAGGGGCAAGTCCTACGTCCAGT  
GCCAGGGGATCCCGCAGGGCTCCATCCTCTCCACGCTGCTCTGCAGCCT  
GTGCTACGGCGACATGGAGAACAAGCTGTTTGCGGGGATTTCGGCGGGAC  
GGGCTGCTCCTGCGTTTGGTGGATGATTTCTTGTTGGTGACACCTCACC  
TCACCCACGCGAAAACCTTCCTCAGGACCCTGGTCCGAGGTGTCCCTGA  
GTATGGCTGCGTGGTGAACCTTGCGGAAGACAGTGGTGAACCTTCCTGTA



**FIGURE 5D**

GAAGACGAGGCCCTGGGTGGCACGGCTTTTGTTCAGATGCCGGCCCACG  
GCCTAT

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

# FIGURE 6A

HASGQRCVLLRTWEALAPATPAMPRAPRCRAVRSLLRSHYREVLPLATF  
 VRRLGPGQWRLVQRGDPAAFRALVAQCLVCVPWDARPPPAAPSFRQVSC  
 LKELVARVLQRLCERGAKNVLAFGFALLDGARGGPPEAFTTSVRSYLPN  
 TVTDALRGSGAWGLLLRRVGDDVLVHLLARCALFVLVAPSCAYQVCGPP  
 LYQLGAATQARPPPHASGPRRRLGCERAWNHSVREAGVPLGLPAPGARR  
 RGGSASRSLPLPKRPRRGAAPEPERTPVGQGSWAHPGRTRGPSDRGFCV  
 VSPARPAEEATSLEGALSGTRHSHPSVGRQHHAGPPSTSRPPRPWDTPC  
 PPVYAETKHFLYSSGDKEQLRPSFLLSSLRPSLTGARRLVETIFLGSRP  
 WMPGTPRRLPRLPQRYWQMRPLFLELLGNHAQCPYGVLLKTHCPLRAAV  
 TPAAGVCAREKPGQGSVAAPEEEDTDPRRLVQLLRQHSSPWQVYGFVRAC  
 LRRLVPPGLWGSRHNERFLRNTKKFISLGKHAKLSLQELTWKMSVRDC  
 AWLRRSPGVGCVPAAEHRLREEILAKFLHWLMSVYVVELLRSFFYVTET  
 TFQKNRLFYRKSVWSKLQSIGIRQHLKRVQLRELSEAEVRQHREARPA  
 LLTSRLRFIPKPDGLRPINMDYVVGARTFRREKRAERLTSRVKALFSV  
 LNYERARRPGLLGASVLGLDDIHRAWRTFVLRVRAQDPPPELYFVKVDV  
 TGAYDTIPQDRLTEVIASIIKPQNTYCVRRYAVVQKAAHGHVRKAFKSH  
 VSTLTDLQPYMRQFVAHLQETSPLRDAVVIEQSSSLNEASSGLFDVFLR  
 FMCHHAVRIRGKSYVQCQGIPQGSILSTLLCSLCYGD MENKLFAGIRRD

**FIGURE 6B**

GLLLRLVDDFLLVTPHLTHAKTFLRTLVRGVPEYGCVVNLRKTVVNF PV

EDEALGGTAFVQMPAHGL

Accession: E047560

# FIGURE 7

TCCCCCTGGTGCGGCCTGCTGCTGGATACCCGGACCCTGGAGGTGCAGAGCGACT  
ACTCCAGCTATGCCCCGGACCTCCATCAGAGCCAGTCTCACCTTCAACCGCGGCT  
TCAAGGCTGGGAGGAACATGCGTCGCAAACCTCTTTGGGGTCTTGCGGCTGAAGT  
GTCACAGCCTGTTTCTGGATTTGCAGGTGAACAGCCTCCAGACGGTGTGCACCA  
ACATCTACAAGATCCTCCTGCTGCAGGCGTACAGGTTTCACGCATGTGTGCTGC  
AGCTCCCATTTTCATCAGCAAGTTTGGAAGAACCCACATTTTTCTGCGCGTCA  
TCTCTGACACGGCCTCCCTCTGCTACTCCATCCTGAAAGCCAAGAACGCAGGGA  
TGTCGCTGGGGGCCAAGGGCGCCGCGGCCCTCTGCCCTCCGAGGCCGTGCAGT  
GGCTGTGCCACCAAGCATTCCTGCTCAAGCTGACTCGACACCGTGTACCTACG  
TGCCACTCCTGGGGTCACTCAGGACAGCCCAGACGCAGCTGAGTCGGAAGCTCC  
CGGGGACGACGCTGACTGCCCTGGAGGCCGCAGCCAACCCGGCACTGCCCTCAG  
ACTTCAAGACCATCCTGGACTGATGGCCACCCGCCCACAGCCAGGCCGAGAGCA  
GACACCAGCAGCCCTGTACGCCGGGCTCTACGTCCCAGGGAGGGAGGGGCGGC  
CCACACCCAGGCCCGCACCGCTGGGAGTCTGAGGCCTGAGTGAGTGTTTGGCCG  
AGGCCTGCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAGGCCTGAGCGAGTGT  
CCAGCCAAGGGCTGAGTGTCCAGCACACCTGCCGTCTTCACTTCCCCACAGGCT  
GGCGCTCGGCTCCACCCAGGGCCAGCTTTTCCTCACCAGGAGCCCGGCTTCCA  
CTCCCCACATAGGAATAGTCCATCCCCTGAT

**FIGURE 8A**

CCACGCGTCCGGGCAGCGCTGCGTCCTGCTGCGCACGTGGGAAGCCCTGGCCCC  
GGCCACCCCCGCGATGCCGCGCGCTCCCCGCTGCCGAGCCGTGCGCTCCCTGCT  
GCGCAGCCACTACCGCGAGGTGCTGCCGCTGGCCACGTTTCGTGCGGCGCCTGGG  
GCCCCAGGGCTGGCGGCTGGTGCAGCGCGGGGACCCGGCGGCTTTCCGCGCGCT  
GGTGGCCCAGTGCCCTGGTGTGCGTGCCCTGGGACGCACGGCCGCCCCCGCCGC  
CCCCTCCTTCCGCCAGGTGTCCTGCCTGAAGGAGCTGGTGGCCCGAGTGCTGCA  
GAGGCTGTGCGAGCGCGGCGCGAAGAACGTGCTGGCCTTCGGCTTCGCGCTGCT  
GGACGGGGCCCCGCGGGGGCCCCCCCCGAGGCCTTCACCACCAGCGTGCGCAGCTA  
CCTGCCCAACACGGTGACCGACGCACTGCGGGGGAGCGGGGCGTGGGGGCTGCT  
GCTGCGCCGCGTGCGGCGACGACGTGCTGGTTCACCTGCTGGCACGCTGCGCGCT  
CTTTGTGCTGGTGGCTCCCAGCTGCGCCTACCAGGTGTGCGGGCCGCGCTGTA  
CCAGCTCGGCGCTGCCACTCAGGCCCCGGCCCCCGCCACACGCTAGTGGACCCCG  
AAGGCGTCTGGGATGCGAACGGGCCTGGAACCATAGCGTCAGGGAGGCCGGGGT  
CCCCCTGGGCCTGCCAGCCCCGGGTGCGAGGAGGCGCGGGGGCAGTGCCAGCCG  
AAGTCTGCCGTTGCCCAAGAGGCCCAGGCGTGCGCTGCCCTGAGCCGGAGCG  
GACGCCCCGTTGGGCAGGGGTCTTGGGCCCACCCGGGCAGGACGCGTGGACCGAG  
TGACCGTGGGTTTCTGTGTGGTGTACCTGCCAGACCCGCCGAAGAAGCCACCTC  
TTTGAGGGTGCGCTCTCTGGCACGCGCCACTCCCACCCATCCGTGGGCCGCCA  
GCACCACGCGGGCCCCCATCCACATCGCGGCCACCACGTCCCTGGGACACGCC  
TTGTCCCCCGGTGTACGCCGAGACCAAGCACTTCCTCTACTCCTCAGGCGACAA

03647069

**FIGURE 8B**

GGAGCAGCTGCGGCCCTCCTTCCTACTCAGCTCTCTGAGGCCAGCCTGACTGG  
CGCTCGGAGGCTCGTGGAGACCATCTTTCTGGGTTCAGGCCCTGGATGCCAGG  
GACTCCCCGCAGGTTGCCCCGCCTGCCCCAGCGCTACTGGCAAATGCGGCCCT  
GTTTCTGGAGCTGCTTGGGAACCACGCGCAGTGCCCCCTACGGGGTGCTCCTCAA  
GACGCACTGCCCCGCTGCGAGCTGCGGTACCCCCAGCAGCCGGTGTCTGTGCCCG  
GGAGAAGCCCCAGGGCTCTGTGGCGGCCCCCGAGGAGGAGGACACAGACCCCCG  
TCGCCTGGTGCAGCTGCTCCGCCAGCACAGCAGCCCCCTGGCAGGTGTACGGCTT  
CGTGCGGGCCTGCCTGCGCCGGCTGGTGCCCCCAGGCCTCTGGGGCTCCAGGCA  
CAACGAACGCCGCTTCCTCAGGAACACCAAGAAGTTCATCTCCCTGGGGAAGCA  
TGCCAAGCTCTCGCTGCAGGAGCTGACGTGGAAGATGAGCGTGCGGGACTGCGC  
TTGGCTGCGCAGGAGCCCAGGGGTGGCTGTGTTCGGCCGCAGAGCACCGTCT  
GCGTGAGGAGATCCTGGCCAAGTTCCTGCACTGGCTGATGAGTGTGTACGTCGT  
CGAGCTGCTCAGGTCTTTCTTTTATGTCACGGAGACCACGTTTCAAAGAACAG  
GCTCTTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTGCAAAGCATTGGAATCAG  
ACAGCACTTGAAGAGGGTGCAGCTGCGGGAGCTGTCGGAAGCAGAGGTCAGGCA  
GCATCGGGAAGCCAGGCCCGCCCTGCTGACGTCCAGACTCCGCTTCATCCCCAA  
GCCTGACGGGCTGCGGCCGATTGTGAACATGGACTACGTCGTGGGAGCCAGAAC  
GTTCCGCAGAGAAAAGAGGGCCGAGCGTCTCACCTCGAGGGTGAAGGCACTGTT  
CAGCGTGCTCAACTACGAGCGGGCGCGCGCCCCGGCCTCCTGGGCGCCTCTGT  
GCTGGGCCTGGACGATATCCACAGGGCCTGGCGCACCTTCGTGCTGCGTGTCG

**FIGURE 8C**

GGCCCAGGACCCGCCGCCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGGCGC  
GTACGACACCATCCCCAGGACAGGCTCACGGAGGTCATCGCCAGCATCATCAA  
ACCCCAGAACACGTACTGCGTGCGTCGGTATGCCGTGGTCCAGAAGGCCGCCCA  
TGGGCACGTCCGCAAGGCCTTCAAGAGCCACGTCTCTACCTTGACAGACCTCCA  
GCCGTACATGCGACAGTTCGTGGCTCACCTGCAGGAGACCAGCCCGCTGAGGGA  
TGCCGTCGTCATCGAGCAGAGCTCCTCCCTGAATGAGGCCAGCAGTGGCCTCTT  
CGACGTCTTCTACGCTTCATGTGCCACCACGCCGTGCGCATCAGGGGCAAGTC  
CTACGTCCAGTGCCAGGGGATCCCGCAGGGCTCCATCCTCTCCACGCTGCTCTG  
CAGCCTGTGCTACGGCGACATGGAGAACAAGCTGTTTGCGGGGATTCGGCGGGA  
CGGGCTGCTCCTGCGTTTGGTGGATGATTTCTTGTTGGTGACACCTCACCTCAC  
CCACGCGAAAACCTTCCTCAGGACCCTGGTCCGAGGTGTCCTTGAGTATGGCTG  
CGTGGTGAAC TTGCGGAAGACAGTGGTGAAC TTCCCTGTAGAAGACGAGGCCCT  
GGGTGGCACGGCTTTTGTT CAGATGCCGGCCACGGCCTATTCCCCTGGTGCGG  
CCTGCTGCTGGATACCCGGACCCTGGAGGTGCAGAGCGACTACTCCAGCTATGC  
CCGGACCTCCATCAGAGCCAGTCTCACCTTCAACCGCGGCTTCAAGGCTGGGAG  
GAACATGCGTCGCAAAC TCTTTGGGGTCTTGCGGCTGAAGTGTCACAGCCTGTT  
TCTGGATTTGCAGGTGAACAGCCTCCAGACGGTGTGCACCAACATCTACAAGAT  
CCTCCTGCTGCAGGCGTACAGGTTTCACGCATGTGTGCTGCAGCTCCCATTTCA  
TCAGCAAGTTTGGAAGAACCCACATTTTTCTGCGCGTCATCTCTGACACGGC  
CTCCCTCTGCTACTCCATCCTGAAAGCCAAGAACGCAGGGATGTCGCTGGGGGC

00943-069  
46907-069

**FIGURE 8D**

CAAGGGCGCCGCCGGCCCTCTGCCCTCCGAGGCCGTGCAGTGGCTGTGCCACCA  
AGCATTCCTGCTCAAGCTGACTCGACACCGTGTCACCTACGTGCCACTCCTGGG  
GTCACTCAGGACAGCCCAGACGCAGCTGAGTCGGAAGCTCCCGGGGACGACGCT  
GACTGCCCTGGAGGCCGCAGCCAACCCGGCACTGCCCTCAGACTTCAAGACCAT  
CCTGGACTGATGGCCACCCGCCACAGCCAGGCCGAGAGCAGACACCAGCAGCC  
CTGTCACGCCGGGCTCTACGTCCCAGGGAGGGAGGGGCGGCCCACACCCAGGCC  
CGCACCGCTGGGAGTCTGAGGCCTGAGTGAGTGTTTGGCCGAGGCCTGCATGTC  
CGGCTGAAGGCTGAGTGTCCGGCTGAGGCCTGAGCGAGTGTCCAGCCAAGGGCT  
GAGTGTCCAGCACACCTGCCGTCTTCACTTCCCCACAGGCTGGCGCTCGGCTCC  
ACCCCAGGGCCAGCTTTTCCTCACCAGGAGCCCGGCTTCCACTCCCCACATAGG  
AATAGTCCATCCCCTGAT

CGGCTGAAGGCTGAGTGTCCGGCTGAGGCCTGAGCGAGTGTCCAGCCAAGGGCT



**FIGURE 9A**

HASGQRCVLLRTWEALAPATPAMPRAPRCRAVRSLLRSHYREVLPLATF  
VRRLGPPQGWRLVQRGDPAAFRALVAQCLVCVPWDARPPPAAPSFRQVSC  
LKELVARVLQRLCERGAKNVLAFGFALLDGARGGPPEAFTTSVRSYLPN  
TVTDALRGSGAWGLLLRRVGDDVLVHLLARCALFVLVAPSCAYQVCGPP  
LYQLGAATQARPPPHASGPRRRLGCERAWNHSVREAGVPLGLPAPGARR  
RGGSASRSLPLPKRPRRGAAPEPERTPVGQGSWAHPGRTRGPSDRGFCV  
VSPARPAEEATSLEGALSGTRHSHPSVGRQHHAGPPSTSRPPRPWDTPC  
PPVYAETKHFLYSSGDKEQLRPSFLLSSLRPSLTGARRLVETIFLGSRP  
WMPGTPRRLPRLPQRYWQMRPLFLELLGNHAQCPYGVLLKTHCPLRAAV  
TPAAGVCAREKPKQGSVAAPEEEDTDPRRLVQLLRQHSSPWQVYGFVRAC  
LRRLVPPGLWGSRHNERFLRNTKKFISLGKHAKLSLQELTWKMSVRDC  
AWLRRSPGVGCVPAAEHRLREEILAKFLHWLMSVYVELLRSFFYVTET  
TFQKNRLFFYRKSVWSKLQSIGIRQHLKRVQLRELSEAEVRQHREARPA  
LLTSRLRFIPKPDGLRPVNM DYVVGARTFRREKRAERLTSRVKALFSV  
LNYERARRPGLLGASVLGLDDIHRAWRTFVLRVRAQDPPPELYFVKVDV  
TGAYDTIPQDRLTEVIASIIKPQNTYCVRRYAVVQKAAHGHVRKAFKSH  
VSTLTDLQPYMRQFVAHLQETSPLRDAVVIEQSSSLNEASSGLFDVFLR  
FMCHHAVRIRGKSYVQCQGIPOGSILSTLLCSLCYGD MENKLFAGIRRD  
GLLLRLVDDFLLVTPHLTHAKTFLRTLVRGVPEYGCVVNLRKTVVNFVPV

**FIGURE 9B**

EDEALGGTAFVQMPAHGLFPWCGLLLDTRTLEVQSDYSSYARTSIRASL  
TFNRGFKAGRNMRRKLFGLRLKCHSLFLDLQVNSLQTVCTNIYKILL  
QAYRFHACVLQLPFHQVWKNPTFFLRVISDTASLCYSILKAKNAGMSL  
GAKGAAGPLPSEAVQWLCHQAFLLKLTRHRVTYVPLLGSLRTAQTQLSR  
KLPGTTLTALEAAANPALPSDFKTILD

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIG. 10

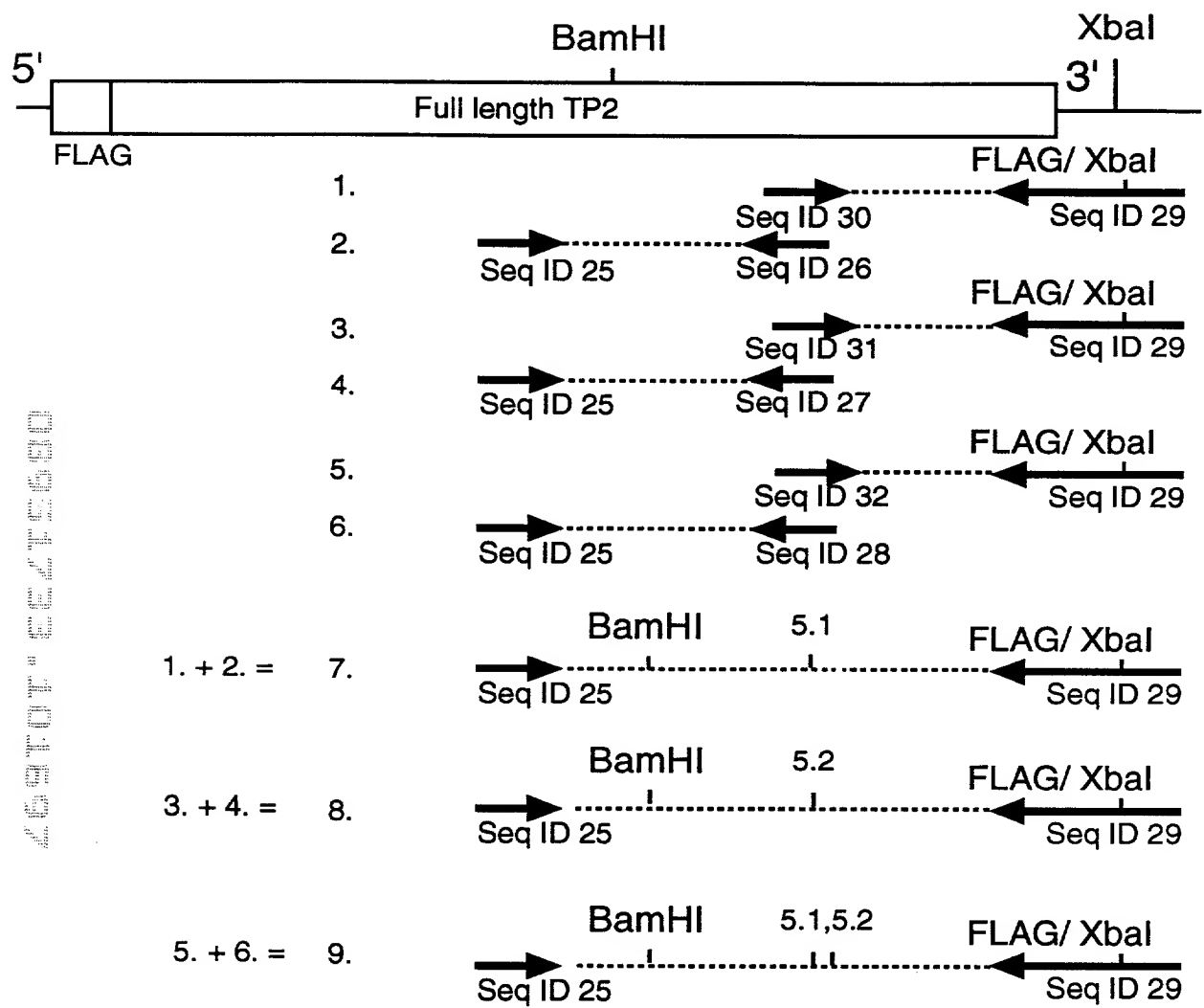


FIG. 14

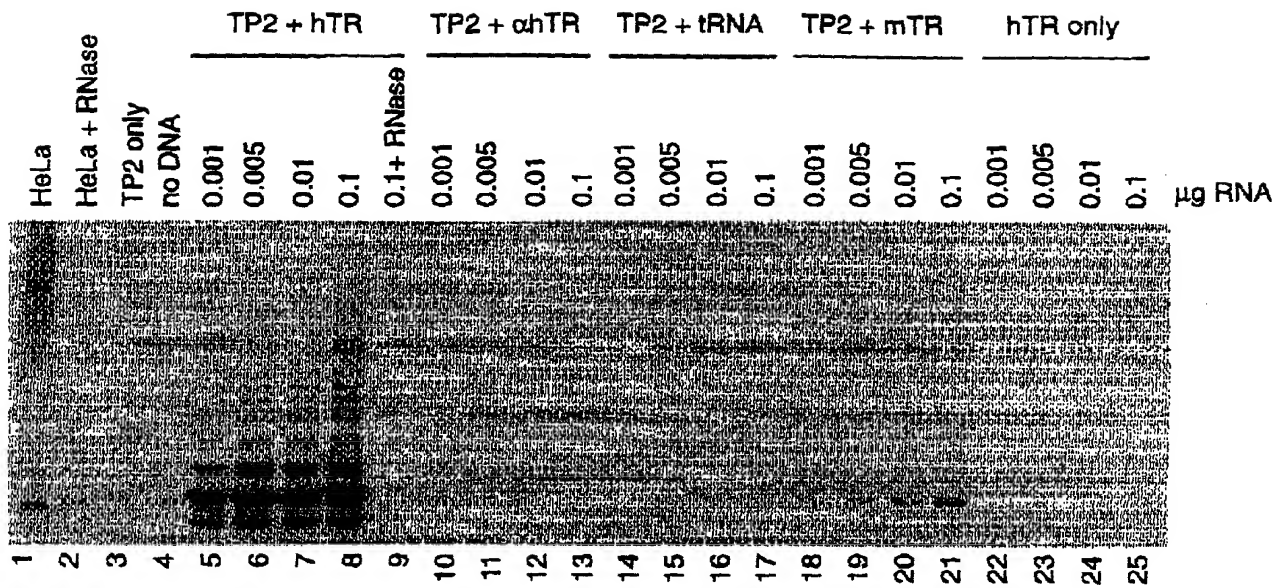
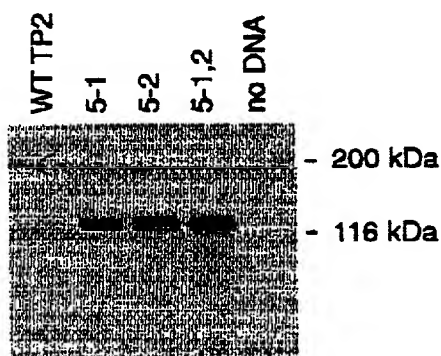
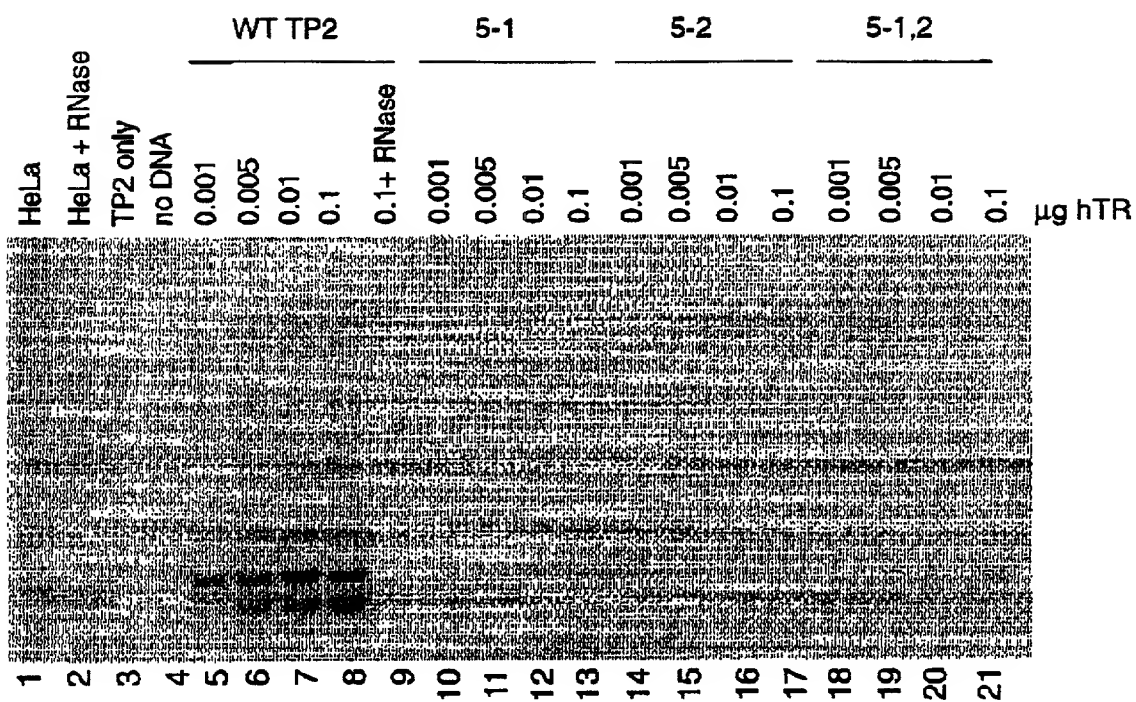


FIG. 15



**FIG. 17**

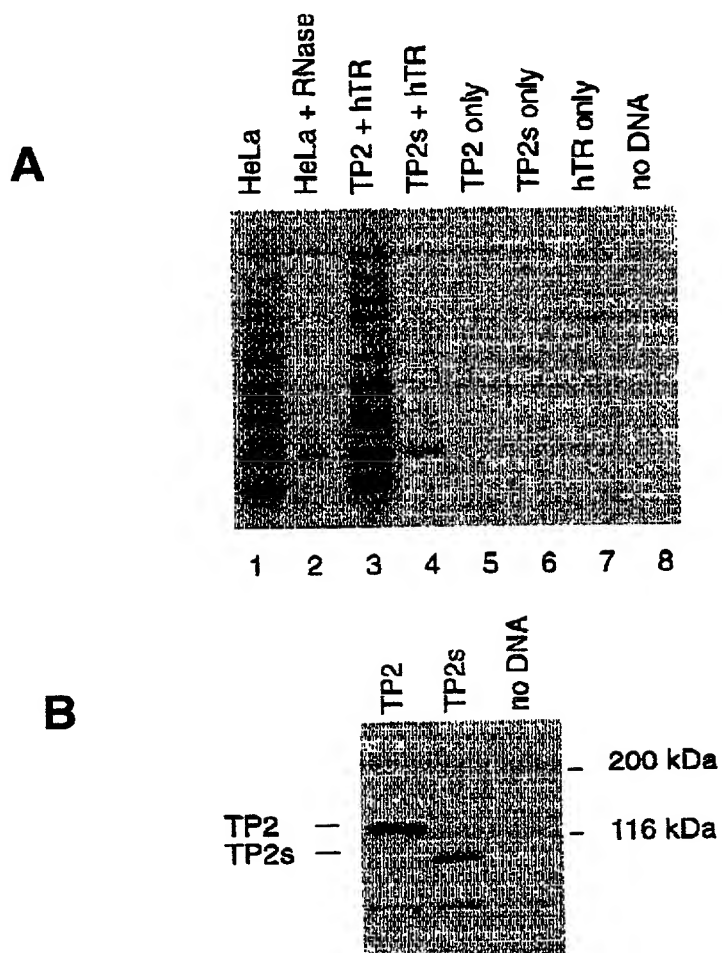
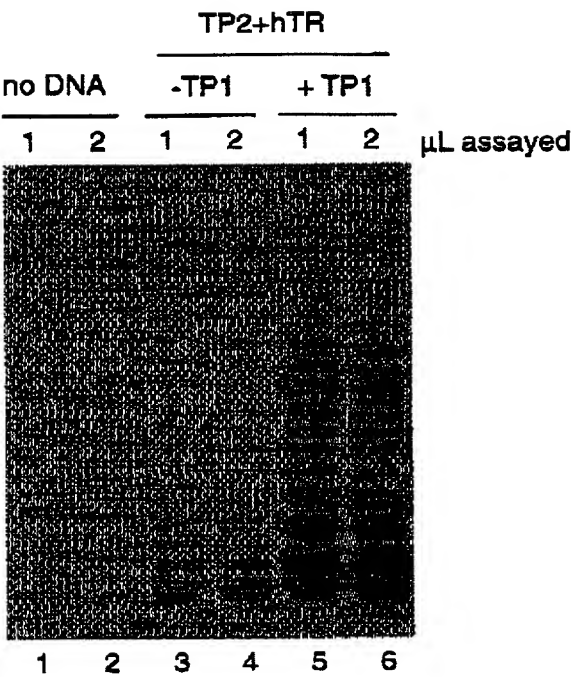
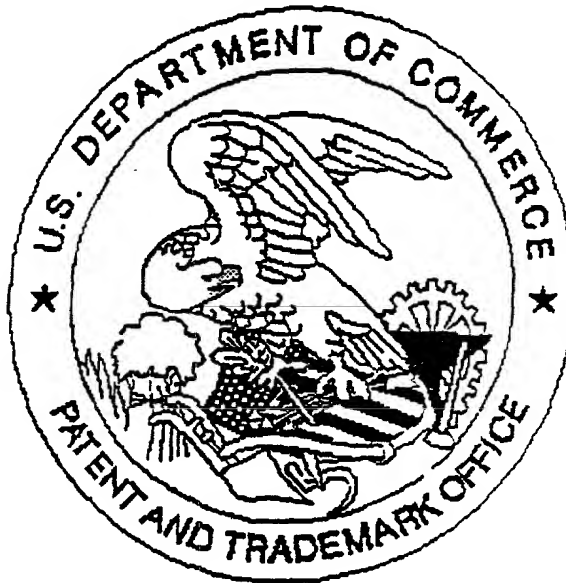


FIG. 18



# United States Patent & Trademark Office

Office of Initial Patent Examination – Scanning Division



## Application deficiencies found during scanning:

1. Application papers are not suitable for scanning and are not in compliance with 37 CFR 1.52 because:
  - ☐ All sheets must be the same size and either A4 (21 cm x 29.7 cm) or 8-1/2" x 11". Pages \_\_\_\_\_ do not meet these requirements.
  - ☐ Papers are not flexible, strong, smooth, non-shiny, durable, and white.
  - ☐ Papers are not typewritten or mechanically printed in permanent ink on one side.
  - ☐ Papers contain improper margins. Each sheet must have a left margin of at least 2.5 cm (1") and top, bottom and right margins of at least 2.0 cm (3/4").
  - ☐ Papers contain hand lettering.
2. Drawings are not in compliance and were not scanned because:
  - ☐ The drawings or copy of drawings are not suitable for electronic reproduction.
  - ☐ All drawings sheets are not the same size. Pages must be either A4 (21 cm x 29.7 cm) or 8-1/2" x 11".
  - ☐ Each sheet must include a top and left margin of at least 2.5 cm (1"), a right margin of at least 1.5 cm (9/16") and a bottom margin of at least 1.0 cm (3/8").
3. Page(s) Fig 11, 12, 13, 16 are not of sufficient clarity, contrast and quality for electronic reproduction.
4. Page(s) \_\_\_\_\_ are missing.
5. OTHER: \_\_\_\_\_

08951733-101697